

see for hjaice54 - alone

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2005, 08:29:12 ; Search time 164 Seconds

(without alignments)

313.653 Million cell updates/sec

Title: HJACE54

Perfect score: 682

Sequence: 1 msrplvpcshalpglspg.....leqlrelriegsvqlcyvhs 133

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

A\_Geneseq\_16Dec04:\*

1: Geneseq1980s:\*

2: Geneseq1990s:\*

3: Geneseq2000s:\*

4: Geneseq2001s:\*

5: Geneseq2002s:\*

6: Geneseq2003as:\*

7: Geneseq2003bs:\*

8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	671	98.4	133	2 AAW61627 Clone HJA
2	671	98.4	133	3 AAY44510 Human Gal
3	671	98.4	133	3 AAB35705 Human Gal
4	671	98.4	133	7 AAW01509 Human rec
5	671	98.4	133	7 AAG40247 Human Gal
6	671	98.4	275	3 AAB35706 Human Gal
7	671	98.4	275	7 ADG40270 Human Gal
8	671	98.4	296	3 AAB35707 Human Gal
9	671	98.4	296	7 ADG40272 Human Gal
10	671	98.4	314	5 AAM48755 Human adi
11	671	98.4	336	3 AAY70029 Human ext
12	124.5	18.3	318	8 ADI13483 Human gal
13	120.5	17.7	171	3 AAY79505 Human car
14	120.5	17.7	196	4 AAB70086 Human sec
15	119	17.4	144	7 ADE57164 Rat Prote
16	119	17.4	144	7 ADE57162 Rat Prote
17	119	17.4	145	7 ADE57162 Protein #
18	119	17.4	145	7 ADG40248 Rat galec
19	118.5	17.4	172	8 ADN04944 Antipsori
20	109.5	16.1	158	6 ABR69609 Human CGD
21	109.5	16.1	258	8 ADO21123 Human car
22	109.5	16.1	258	8 ABR83347 Human dia
23	109.5	16.1	258	8 ABR83348 Human dia
24	109.5	16.1	316	3 AAY87403 Human PCT
25	109.5	16.1	316	4 AAB85030 Protein e

26	109.5	16.1	316	8 ADO57924 Human gal
27	109.5	16.1	317	8 AAW56505 Human gal
28	109.5	16.1	317	8 ADS34896 Human aut
29	109.5	16.1	317	8 ADS34895 Human aut
30	109.5	16.1	318	5 ABR65200 Hypoxia-r
31	109.5	16.1	318	7 ABO1523 Protein #
32	109.5	16.1	318	7 ADG40249 Human gal
33	109.5	16.1	318	7 ADG91609 Human lec
34	109.5	16.1	318	8 ADS88201 Human pro
35	108.5	15.9	316	6 ABR47558 Breast ca
36	108.5	15.9	317	2 AAW03519 Prostate
37	108.5	15.9	317	6 ABO62772 Human pro
38	106.5	15.6	368	3 AAY87405 Human PCT
39	106.5	15.6	368	4 AAB5032 Protein e
40	102	15.0	322	7 ADC53842 Mouse gal
41	98	14.4	322	7 ABO63650 Rat urate
42	98	14.4	354	7 ADE62927 Rat Prote
43	98	14.4	354	7 ADE62927 Rat Prote
44	98	14.4	354	7 ADD48099 Human aut
45	97.5	14.3	359	8 ADS34894 Human aut
46	96.5	14.1	218	8 ADO57926 Human gal
47	96.5	14.1	358	3 AAY87404 Human PCT
48	96.5	14.1	358	4 AAB85031 Protein e
49	96.5	14.1	359	7 ADG91610 Human lec
50	96.5	14.1	359	8 ADN04357 Antipsori
51	96.5	14.1	359	8 ADO28663 Human gal
52	96.5	14.1	359	8 ADP54781 Human PRO
53	96	14.1	329	5 AAY97036 Human bla
54	95.5	14.0	125	4 AAB70068 Human sec
55	95.5	14.0	125	5 ABR65507 Human alb
56	95.5	14.0	125	5 ABR43710 Human car
57	95.5	14.0	125	5 ADL78774 Albumin f
58	95.5	14.0	131	6 ABO62773 Second ca
59	95.5	14.0	135	8 ADJ66636 Gallecin-
60	95.5	14.0	136	2 AAY23992 Amino aci
61	95.5	14.0	136	5 AAY97823 Human cel
62	95.5	14.0	136	7 ADF30697 Human ang
63	95.5	14.0	136	8 ADN17281 Human gal
64	95.5	14.0	136	8 ABR82094 Tumour-as
65	93	13.6	262	8 ADN17284 Chicken g
66	91.5	13.4	316	3 AAY87406 Murine PC
67	91.5	13.4	359	8 ADN04496 Antipsori
68	91.5	13.4	359	8 ADP54924 Human PRO
69	90.5	13.3	139	3 AAG02648 Human sec
70	88.5	13.0	264	2 AAW71219 Beta-D-g
71	88.5	13.0	264	6 ABO4782 Human exp
72	87.5	12.8	300	5 AAY97824 Human cel
73	87.5	12.8	316	2 AAR75702 Rat galec
74	87.5	12.8	316	7 ADD90598 Rat galec
75	87.5	12.8	316	8 ADI13481 Rat galec
76	87	12.8	140	4 ABO22960 Novel hum
77	87	12.8	355	6 ABO69163 Human NOV
78	86.5	12.7	264	2 AAW88106 Human gal
79	86.5	12.7	264	4 AAG67687 Amino aci
80	86.5	12.7	333	5 AAD90724 Ornithodo
81	85.5	12.5	263	5 AAY97822 Mouse cel
82	85.5	12.5	264	2 AAR12531 Mac2.16 e
83	85.5	12.5	264	7 ADC53841 Mouse gal
84	85.5	12.5	264	8 ADP74020 Murine ga
85	85.5	12.5	278	2 AAR13338 Mac-2 pro
86	84	12.3	241	4 AAG92256 C glutami
87	83	12.2	355	8 ADO08329 Human NOV
88	82.5	12.1	136	8 ADN17287 Mouse gal
89	82.5	12.1	136	8 ADN17286 Rat galec
90	82.5	12.1	250	2 AAW71218 Beta-D-ga
91	82.5	12.1	250	6 ABO4684 Human exp
92	81.5	12.0	268	4 ABR75625 Human col
93	81.5	12.0	293	8 ABR84544 Human dia
94	81.5	12.0	323	2 AAW06551 Human col
95	81.5	12.0	323	2 AAW11841 Human gal
96	81.5	12.0	323	2 AAW56503 Human gal
97	81.5	12.0	323	2 AAW46883 Protein s
98	81.5	12.0	323	4 AAM24512 C892P pre

99 81.5 12.0 323 5 AA015235 Human gal  
100 81.5 12.0 323 5 AAU76211 Human col

ALIGNMENTS

RESULT 1

AAW61627

ID AAW61627 standard; protein; 133 AA.

XX AC AAW61627;

XX DT 27-OCT-1998 (first entry)

XX DE Clone HUACE54 of Lectin superfamily.

XX KW Human; receptor; immune disorder; cancers; blood disorder;

XX KW Juvenile rheumatoid arthritis; Graves disease.

XX OS Homo sapiens.

XX PN WO9831799-A2.

XX PD 23-JUL-1998.

XX PF 21-JAN-1998; 98WO-US000959.

XX PR 21-JAN-1997; 97US-0034204P.

XX PR 21-JAN-1997; 97US-0034205P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Ni J, Gentz RL, Rosen CA;

XX WPI; 1998-427559/36.

XX DR N-PSDB; AAW48123.

XX PT New isolated polynucleotide(s) and encoded receptor polypeptide(s) -

XX PT used to develop products for diagnosing or treating e.g. immune

XX PT disorders, cancers, blood disorders or immuno-compromised disease states.

XX PS Claim 11; Page 53; 79pp; English.

XX CC Clone HUACE54 is a member of the Lectin receptor superfamily. The

XX CC products generated using the receptor can be used for treating abnormal

XX CC conditions related to both an excess of and insufficient amounts of

XX CC receptor activity. They can be used in the treatment of e.g. immune

XX CC disorders, cancers, blood disorders, juvenile rheumatoid arthritis,

XX CC Graves disease or immunocompromised disease states. The products can also

XX CC be used for detection and diagnosis

XX SQ Sequence 133 AA;

Query Match 98.4%; Score 671; DB 2; Length 133;

Best Local Similarity 99.2%; Pred. No. 6e-76;

Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSPRLEVPCHALPQGLSPGQVIIVRGLVLPQPKFTVSLRDOAAHAPVTLRASFDRTL 60

DB 1 MSPRLEVPCHALPQGLSPGQVIIVRGLVLPQPKFTVSLRDOAAHAPVTLRASFDRTL 60

QY 61 AWISRWGQKLLSAPFLFPQRFPEVLLLFQEGGLKALNGQGLGATSMNQALEQRL 120

DB 61 AWISRWGQKLLSAPFLFPQRFPEVLLLFQEGGLKALNGQGLGATSMNQALEQRL 120

QY 121 RISGSVQLYCVHS 133

DB 121 RISGSVQLYCVHS 133

RESULT 2

AAW44510

AAW44510

AAW44510 standard; protein; 133 AA.

AAW44510;

DT 04-APR-2000 (first entry)

DE Human Galectin 11 protein.

KW Galectin 11; HUACE54 plasmid; galectin 5; galectin 3; galectin 8;

KW cell differentiation; proliferation; metastasis; tumour; apoptosis;

KW anti-galectin 11 antibody; treatment; autoimmune disease; cancer;

KW inflammatory disease; asthma; allergy; neuropathy.

OS Homo sapiens.

FH Key Location/Qualifiers

FT Domain 65..70

FT Domain /label= Epitope

FT Domain 118..124

FT Domain /label= Epitope

PN WO200001728-A1.

XX PD 13-JAN-2000.

XX PF 02-JUL-1999; 99WO-US015169.

XX PR 06-JUL-1998; 98US-00109864.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Rosen CA, Ni J, Gentz RL;

XX WPI; 2000-126931/11.

XX DR N-PSDB; AAZ49417.

XX PT New human galectin polypeptide, useful for treating cell growth disorders

XX PT in mammals e.g. cancer, autoimmune diseases, inflammatory diseases,

XX PT asthma and allergic diseases.

XX PS Claim 1; Fig 1; 99pp; English.

XX CC The present sequence is the human galectin 11 protein. It is encoded by

XX CC cDNA derived from HUACE54 plasmid (ATCC 209053). Galectin 11 shows

XX CC homology to rat galectin 5, chicken galectin 3 and human galectin 8. It

XX CC is involved in cell differentiation, proliferation, metastasis of tumour

XX CC cells, cell adhesion, autoimmunity, inflammation, cell-cell substrate

XX CC interactions and apoptosis. The polynucleotides can be used to detect

XX CC cell growth/differentiation disorders and to produce probes and primers

XX CC for detecting polynucleotides and gene mapping. Anti-galectin 11

XX CC antibodies can be used therapeutically and for purification of proteins.

XX CC Galectin 11 can be used to treat cancer, autoimmune diseases,

XX CC inflammatory diseases, asthma, allergy and neuropathies

XX SQ Sequence 133 AA;

Query Match 98.4%; Score 671; DB 3; Length 133;

Best Local Similarity 99.2%; Pred. No. 6e-76;

Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSPRLEVPCHALPQGLSPGQVIIVRGLVLPQPKFTVSLRDOAAHAPVTLRASFDRTL 60

DB 1 MSPRLEVPCHALPQGLSPGQVIIVRGLVLPQPKFTVSLRDOAAHAPVTLRASFDRTL 60

QY 61 AWISRWGQKLLSAPFLFPQRFPEVLLLFQEGGLKALNGQGLGATSMNQALEQRL 120

DB 61 AWISRWGQKLLSAPFLFPQRFPEVLLLFQEGGLKALNGQGLGATSMNQALEQRL 120

QY 121 RISGSVQLYCVHS 133

DB 121 RISGSVQLYCVHS 133





QY 1 MSPRLVPCSHALPQGLSPGVIIIRGLVLQEPKHTVSLRDOAAHAPVTLRASFAFRTL 60  
 |||||  
 Db 1 MSPRLVPCSHALPQGLSPGVIIIRGLVLQEPKHTVSLRDOAAHAPVTLRASFAFRTL 60  
 |||||  
 QY 61 AMISRWGQKKLISAPFLFYPPQRFVLLLFQEGGLKIALNGQGLGATSMNQALEQUREL 120  
 |||||  
 Db 61 AMISRWGQKKLISAPFLFYPPQRFVLLLFQEGGLKIALNGQGLGATSMNQALEQUREL 120  
 |||||  
 QY 121 RISGSVQLYCVHS 133  
 |||||  
 Db 121 RISGSVQLYCVHS 133  
 |||||  
 RESULT 5  
 ADG40247  
 ID ADG40247 standard; protein; 133 AA.  
 XX  
 AC ADG40247;  
 XX  
 DT 26-FEB-2004 (first entry)  
 XX  
 DE Human Galectin 11 protein.  
 XX  
 KW Galectin 11; cytostatic; immunosuppressive; antiinflammatory;  
 KW antiasthmatic; antiallergic; Gene therapy; cell growth disorder; cancer;  
 KW autoimmune disease; inflammatory disease; asthma; allergic disease;  
 KW human.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2003208044-A1.  
 XX  
 PD 06-NOV-2003.  
 XX  
 PF 06-JUN-2003; 2003US-00455366.  
 XX  
 PR 21-JAN-1997; 97US-0034204P.  
 PR 21-JAN-1997; 97US-0034205P.  
 PR 21-JAN-1998; 98US-00010146.  
 PR 06-JUL-1998; 98US-00109864.  
 PR 21-APR-1999; 99US-0130390P.  
 PR 10-DEC-1999; 99US-0169932P.  
 PR 21-APR-2000; 2000US-00557170.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PA (LJOL-) LA JOLLA INST ALLERGY & IMMUNOLOGY.  
 XX  
 PI Ni J, Gentz RL, Rosen CA, Liu F;  
 XX  
 DR WPI: 2003-864797/80.  
 DR N-PSDB; ADG40246.  
 XX  
 PT New galectin 11 polypeptides, useful for treating a cell growth disorder,  
 PT e.g. cancer, autoimmune diseases, inflammatory diseases, asthma, or  
 PT allergic diseases.  
 XX  
 PS Claim 9; SEQ ID NO 2; 123pp; English.  
 XX  
 CC The invention relates to an isolated polynucleotide comprising a sequence  
 CC that is at least 95% identical to a sequence encoding Human Galectin 11,  
 CC galectin 11alpha, galectin 11beta (or their defined fragments). Also  
 CC included are a method of making a recombinant vector by inserting the  
 CC isolated polynucleotide into a vector, a recombinant vector comprising  
 CC the polynucleotide, a genetically engineered host cell comprising the  
 CC polynucleotide, a method of producing a galectin 11 polypeptide (by  
 CC culturing the genetically engineered host cell under conditions suitable  
 CC to produce the polypeptide, and recovering the polypeptide), an isolated  
 CC galectin 11 polypeptide comprising any of the amino acid sequences cited  
 CC above, a pharmaceutical composition comprising the polypeptide and a  
 CC carrier, an isolated antibody that binds specifically to the polypeptide,  
 CC a method of detecting a galectin 11 polypeptide in a sample (by  
 CC contacting the sample with the antibody, and detecting the presence of

CC the antibody bound to the polypeptide), a method of treating a cell  
 CC growth disorder in a mammal by administering the polypeptide to the  
 CC mammal, and a method of regulating cell growth or differentiation in a  
 CC mammal by administering a galectin 11 polypeptide or polynucleotide to  
 CC the mammal to suppress cell growth or differentiation. The  
 CC polynucleotide, polypeptide and composition are useful for treating a  
 CC cell growth disorder, e.g. cancer, autoimmune diseases, inflammatory  
 CC diseases, asthma, or allergic diseases (many examples cited in the  
 CC specification). The present sequence represents a galectin 11 protein.  
 XX  
 SQ Sequence 133 AA;  
 Query Match 98.4%; Score 671; DB 7; Length 133;  
 Best Local Similarity 99.2%; Pred. No. 6e-76;  
 Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MSPRLVPCSHALPQGLSPGVIIIRGLVLQEPKHTVSLRDOAAHAPVTLRASFAFRTL 60  
 |||||  
 Db 1 MSPRLVPCSHALPQGLSPGVIIIRGLVLQEPKHTVSLRDOAAHAPVTLRASFAFRTL 60  
 |||||  
 QY 61 AMISRWGQKKLISAPFLFYPPQRFVLLLFQEGGLKIALNGQGLGATSMNQALEQUREL 120  
 |||||  
 Db 61 AMISRWGQKKLISAPFLFYPPQRFVLLLFQEGGLKIALNGQGLGATSMNQALEQUREL 120  
 |||||  
 QY 121 RISGSVQLYCVHS 133  
 |||||  
 Db 121 RISGSVQLYCVHS 133  
 |||||  
 RESULT 6  
 AAB35706  
 ID AAB35706 standard; protein; 275 AA.  
 XX  
 AC AAB35706;  
 XX  
 DT 15-FEB-2001 (first entry)  
 XX  
 DE Human galectin 11alpha amino acid sequence.  
 XX  
 DE Human; galectin 11; immunosuppressive; thyromimetic; hepatotropic;  
 KW antiinflammatory; vasotropic; dermatological; nephrotropic; cytostatic;  
 KW antirheumatic; antiallergic; antiviral; anti-HIV; nootropic; cardiant;  
 KW neuroprotective; antiparkinsonian; ophthalmological; antianaemic;  
 KW cerebroprotective; antibacterial; immunomodulatory; cancer; allergy;  
 KW autoimmune disease; inflammatory disease; arthritis; Parkinson's disease;  
 KW Alzheimer's disease; ischaemia; stroke; anorexia; shock.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200063221-A2.  
 XX  
 PD 26-OCT-2000.  
 XX  
 PF 21-APR-2000; 2000WO-US010714.  
 XX  
 PR 21-APR-1999; 99US-0130390P.  
 PR 10-DEC-1999; 99US-0169932P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PA (LJOL-) LAJOLLA INST ALLERGY & IMMUNOLOGY.  
 XX  
 PI Ni J, Rosen CA, Gentz RL, Lui F;  
 XX  
 DR WPI: 2000-665238/64.  
 DR N-PSDB; AAC66207.  
 XX  
 PT Novel galectin polynucleotides and polypeptides, used in the diagnosis  
 PT and treatment of cancers, autoimmune disorders, and inflammatory  
 PT disorders, and to screen for antagonists and agonists.  
 XX  
 PS Claim 9; Fig 6; 314pp; English.  
 XX  
 CC Polynucleotides AAC66206 - AAC66208 encode human galectin 11, 11alpha and

CC libeta AAB35705 - AAB35707 respectively. Galectin 11 polynucleotide and  
CC protein sequences have immunosuppressive; thyromimetic; hepatotropic;  
CC antiinflammatory; vasotropic; dermatological; nephrotropic; antirheumatic  
CC ; antiarthritic; antiviral; anti-HIV; nootropic; neuroprotective;  
CC antiparkinsonian; ophthalmological; cytostatic; antianemic; cardiac;  
CC cerebroprotective; antibacterial; and immunomodulatory activity. Galectin  
CC 11 polypeptides can be used to treat cell growth disorders, particularly  
CC cancer, autoimmune diseases, inflammatory diseases, asthma, or allergic  
CC diseases, in a mammal. The proteins and polynucleotides can be  
CC administered to regulate cell growth or differentiation in mammals.  
CC Antibodies specific for galectin 11 can be used to detect galectin 11  
CC polypeptides in a samples, used in diagnosis techniques. Proteins, in  
CC polynucleotides and antibodies, can be used to treat, e.g. Hashimoto's  
CC thyroiditis, biliary cirrhosis, Behcet's disease, Crohn's disease,  
CC polymyositis, systemic lupus erythematosus, glomerulonephritis,  
CC rheumatoid arthritis, viral infections, inflammation, graft versus host  
CC disease, graft rejection, acquired immunodeficiency syndrome, Alzheimer's  
CC disease, Parkinson's disease, amyotrophic lateral sclerosis, Retinitis  
CC pigmentosa, cerebellar degeneration, brain tumour, multiple sclerosis,  
CC Sjogren's syndrome, aplastic anaemia, ischaemia, myocardial infarction,  
CC stroke, reperfusion injury, liver injury, septic shock, cachexia, and  
CC anorexia. PCR primers AAC66211 - AAC66215 are used to amplify galectin 11  
CC encoding DNA. Oligonucleotides AAC66216 - AAC66224 are used in the  
CC identification and characterisation of the galectin 11 polynucleotides  
XX  
SQ Sequence 275 AA;

Query Match 98.4%; Score 671; DB 3; Length 275;  
Best Local Similarity 99.2%; Pred. No. 1.6e-75;  
Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSPRLEVPCHALPQGLSPGQVIIVRGLVLPQEPKFTVSLRDQAAHAPVTLRASFAADRTL 60  
DB 143 MSPRLEVPCHALPQGLSPGQVIIVRGLVLPQEPKFTVSLRDQAAHAPVTLRASFAADRTL 202  
QY 61 AWISRWQKKLISAPFLFYQRFPEVLLLFQEGGLKALNGQGLGATSMNQALEQREL 120  
DB 203 AWISRWQKKLISAPFLFYQRFPEVLLLFQEGGLKALNGQGLGATSMNQALEQREL 262  
QY 121 RIGSVOLYCVHS 133  
DB 263 RIGSVOLYCVHS 275

RESULT 7  
ADG40270  
ID ADG40270 standard; protein; 275 AA.  
XX  
AC ADG40270;  
XX  
XX 26-FEB-2004 (first entry)  
XX  
XX Human Galectin 11alpha protein.  
XX  
XX Galectin 11; cytostatic; immunosuppressive; antiinflammatory;  
XX antiasthmatic; antiallergic; gene therapy; cell growth disorder; cancer;  
XX autoimmune disease; inflammatory disease; asthma; allergic disease;  
XX human.  
XX  
XX Homo sapiens.  
XX  
XX  
XX Key Location/Qualifiers  
XX 1. 121  
XX Region /note= "Claimed in claim 9"  
XX FT 151. .275  
XX Region /note= "Claimed in claim 9"  
XX FT  
XX FT  
XX US2003208044-A1.  
XX PN  
XX  
XX 06-NOV-2003.  
XX PD  
XX 06-JUN-2003; 2003US-00455366.  
XX PF  
XX  
XX

PR 21-JAN-1997; 97US-0034204P.  
PR 21-JAN-1997; 97US-0034205P.  
PR 21-JAN-1998; 98US-00010146.  
PR 06-JUL-1998; 98US-00109864.  
PR 21-APR-1999; 99US-0130390P.  
PR 10-DEC-1999; 99US-0169932P.  
PR 21-APR-2000; 2000US-00557170.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA (LJOL-) LA JOLLA INST ALLERGY & IMMUNOLOGY.  
XX  
XX Ni J, Gentz RL, Rosen CA, Liu F;  
PI WPI; 2003-864797/80.  
XX N-PSDB; ADG40269.  
DR  
XX New galectin 11 polypeptides, useful for treating a cell growth disorder,  
PT e.g. cancer, autoimmune diseases, inflammatory diseases, asthma, or  
PT allergic diseases.  
XX  
XX Claim 9; SEQ ID NO 25; 123pp; English.  
XX  
XX The invention relates to an isolated polynucleotide comprising a sequence  
CC that is at least 95% identical to a sequence encoding Human Galectin 11,  
CC galectin 11alpha, galectin 11beta (or their defined fragments). Also  
CC included are a method of making a recombinant vector by inserting the  
CC isolated polynucleotide into a vector, a recombinant vector comprising  
CC the polynucleotide, a genetically engineered host cell comprising the  
CC polynucleotide, a method of producing a galectin 11 polypeptide (by  
CC culturing the genetically engineered host cell under conditions suitable  
CC to produce the polypeptide, and recovering the polypeptide), an isolated  
CC galectin 11 polypeptide comprising any of the amino acid sequences cited  
CC above, a pharmaceutical composition comprising the polypeptide and a  
CC carrier, an isolated antibody that binds specifically to the polypeptide,  
CC a method of detecting a galectin 11 polypeptide in a sample (by  
CC contacting the sample with the antibody, and detecting the presence of  
CC the antibody bound to the polypeptide), a method of treating a cell  
CC growth disorder in a mammal by administering the polypeptide to the  
CC mammal, and a method of regulating cell growth or differentiation in a  
CC mammal by administering a galectin 11 polypeptide or polynucleotide to  
CC the mammal to suppress cell growth or differentiation. The  
CC polynucleotide, polypeptide and composition are useful for treating a  
CC cell growth disorder, e.g. cancer, autoimmune diseases, inflammatory  
CC diseases, asthma, or allergic diseases (many examples cited in the  
CC specification). The present sequence represents a galectin 11 protein.  
XX  
SQ Sequence 275 AA;

Query Match 98.4%; Score 671; DB 7; Length 275;  
Best Local Similarity 99.2%; Pred. No. 1.6e-75;  
Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSPRLEVPCHALPQGLSPGQVIIVRGLVLPQEPKFTVSLRDQAAHAPVTLRASFAADRTL 60  
DB 143 MSPRLEVPCHALPQGLSPGQVIIVRGLVLPQEPKFTVSLRDQAAHAPVTLRASFAADRTL 202  
QY 61 AWISRWQKKLISAPFLFYQRFPEVLLLFQEGGLKALNGQGLGATSMNQALEQREL 120  
DB 203 AWISRWQKKLISAPFLFYQRFPEVLLLFQEGGLKALNGQGLGATSMNQALEQREL 262  
QY 121 RIGSVOLYCVHS 133  
DB 263 RIGSVOLYCVHS 275

RESULT 8  
AAB35707  
ID AAB35707 standard; protein; 296 AA.  
XX  
XX AAB35707;  
XX  
XX 15-FEB-2001 (first entry)  
XX  
XX

DE Human galectin 11beta amino acid sequence.

XX Human; galectin 11; immunosuppressive; thyromimetic; hepatotropic; antiinflammatory; vasotropic; dermatological; nephrotropic; cytostatic; antirheumatic; antiarthritic; antiviral; anti-HIV; nootropic; cardiant; neuroprotective; antiParkinsonian; ophthalmological; antianaemic; cerebroprotective; antibacterial; immunomodulatory; cancer; allergy; autoimmune disease; inflammatory disease; arthritis; Parkinson's disease; Alzheimer's disease; ischaemia; stroke; anorexia; shock.

XX Homo sapiens.

XX WO200063221-A2.

XX 26-OCT-2000.

XX 21-APR-2000; 2000WO-US010714.

XX 21-APR-1999; 99US-0130390P.

XX 10-DEC-1999; 99US-0169932P.

XX (HUMA-) HUMAN GENOME SCI INC.

PA (LJOL-) LAJOLLA INST ALLERGY & IMMUNOLOGY.

XX Ni J, Rosen CA, Gentz RL, Lui F;

XX WPI; 2000-665238/64.

DR N-PSDB; AAC66208.

XX Novel galectin polynucleotides and polypeptides, used in the diagnosis and treatment of cancers, autoimmune disorders, and inflammatory disorders, and to screen for antagonists and agonists.

XX Claim 9; Fig 7; 314pp; English.

XX Polynucleotides AAC66206 - AAC66208 encode human galectin 11, 11alpha and 11beta AAB35705 - AAB35707 respectively. Galectin 11 polynucleotide and protein sequences have immunosuppressive; thyromimetic; hepatotropic; antiinflammatory; vasotropic; dermatological; nephrotropic; antirheumatic; antiarthritic; antiviral; anti-HIV; nootropic; neuroprotective; antiParkinsonian; ophthalmological; cytostatic; antianaemic; cardiant; cerebroprotective; antibacterial; and immunomodulatory activity. Galectin 11 polypeptides can be used to treat cell growth disorders, particularly cancer, autoimmune diseases, inflammatory diseases, asthma, or allergic diseases, in a mammal. The proteins and polynucleotides can be administered to regulate cell growth or differentiation in mammals.

CC Antibodies specific for galectin 11 can be used to detect galectin 11 polypeptides in a samples, used in diagnosis techniques. Proteins, polynucleotides and antibodies, can be used to treat, e.g. Hashimoto's thyroiditis, biliary cirrhosis, Behcet's disease, Crohn's disease, polynovitis, systemic lupus erythematosus, glomerulonephritis, rheumatoid arthritis, viral infections, inflammation, graft versus host disease, graft rejection, acquired immunodeficiency syndrome, Alzheimer's disease, Parkinson's disease, amyotrophic lateral sclerosis, Retinitis pigmentosa, cerebellar degeneration, brain tumour, multiple sclerosis, Sjogren's syndrome, aplastic anaemia, ischaemia, myocardial infarction, stroke, reperfusion injury, liver injury, septic shock, cachexia, and anorexia. PCR primers AAC66211 - AAC66215 are used to amplify galectin 11 encoding DNA. Oligonucleotides AAC66216 - AAC66224 are used in the identification and characterisation of the galectin 11 polynucleotides

XX Sequence 296 AA;

Query Match 98.4%; Score 671; DB 3; Length 296;

Best Local Similarity 99.2%; Pred. No. 1.8e-75;

Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSRLRVPCHALPQGLSPGVIIIVGLVLOEPKHTVSIROQAAPVTLRASPADRTL 60  
|||||

Db 164 MSRLRVPCHALPQGLSPGVIIIVGLVLOEPKHTVSIROQAAPVTLRASPADRTL 223  
|||||

QY 61 AWISRWGQKKLISAPFLFYPPQRFVFFVLLIFQEGGLKALNGQGLGATSMNQALEQLREL 120  
|||||

Db 224 AWISRWGQKKLISAPFLFYPPQRFVFFVLLIFQEGGLKALNGQGLGATSMNQALEQLREL 283

QY 121 RIGGSVOLYCVHS 133  
|||||

Db 284 RIGGSVOLYCVHS 296  
|||||

RESULT 9

ADG40272

ID ADG40272 standard; protein; 296 AA.

XX AC ADG40272;

XX 26-FEB-2004 (first entry)

DT Human Galectin 11beta protein.

XX Galectin 11; cytostatic; immunosuppressive; antiinflammatory; antiasthmatic; antiallergic; gene therapy; cell growth disorder; cancer; autoimmune disease; inflammatory disease; asthma; allergic disease; human.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Region 1..142

FT /note= "Claimed in claim 9"

XX US2003208044-A1.

XX 06-NOV-2003.

XX 06-JUN-2003; 2003US-00455366.

XX 21-JAN-1997; 97US-0034204P.

XX 21-JAN-1997; 97US-0034205P.

XX 21-JAN-1998; 98US-00010146.

XX 06-JUL-1998; 98US-00109864.

XX 21-APR-1999; 99US-0130390P.

XX 10-DEC-1999; 99US-0169932P.

XX 21-APR-2000; 2000US-00557170.

XX (HUMA-) HUMAN GENOME SCI INC.

PA (LJOL-) LA JOLLA INST ALLERGY & IMMUNOLOGY.

XX Ni J, Gentz RL, Rosen CA, Lui F;

DR WPI; 2003-864797/80.

DR N-PSDB; ADG40271.

XX New galectin 11 polypeptides, useful for treating a cell growth disorder, e.g. cancer, autoimmune diseases, inflammatory diseases, asthma, or allergic diseases.

XX Claim 9; SEQ ID NO 27; 123pp; English.

XX The invention relates to an isolated polynucleotide comprising a sequence that is at least 95% identical to a sequence encoding Human Galectin 11, galectin 11alpha, galectin 11beta (or their defined fragments). Also included are a method of making a recombinant vector by inserting the isolated polynucleotide into a vector, a recombinant vector comprising the polynucleotide, a genetically engineered host cell comprising the polynucleotide, a method of producing a galectin 11 polypeptide (by culturing the genetically engineered host cell under conditions suitable to produce the polypeptide, and recovering the polypeptide), an isolated galectin 11 polypeptide comprising any of the amino acid sequences cited above, a pharmaceutical composition comprising the polypeptide and a carrier, an isolated antibody that binds specifically to the polypeptide, a method of detecting a galectin 11 polypeptide in a sample (by contacting the sample with the antibody, and detecting the presence of the antibody bound to the polypeptide), a method of treating a cell growth disorder in a mammal by administering the polypeptide to the mammal, and a method of regulating cell growth or differentiation in a

CC mammal by administering a galectin 11 polypeptide or polynucleotide to  
CC the mammal to suppress cell growth or differentiation. The  
CC polynucleotide, polypeptide and composition are useful for treating a  
CC cell growth disorder, e.g. cancer, autoimmune diseases, inflammatory  
CC diseases, asthma, or allergic diseases (many examples cited in the  
CC specification). The present sequence represents a galectin 11 protein.  
XX  
SQ Sequence 296 AA;

Query Match 98.4%; Score 671; DB 7; Length 296;  
Best Local Similarity 99.2%; Pred. No. 1.9e-75;  
Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSPRLVPCSHALPQGLSPGQVIIVRGLVLPQPKHPTVSLRDQAAHAPVTLRASFADRTL 60  
Db 164 MSPRLVPCSHALPQGLSPGQVIIVRGLVLPQPKHPTVSLRDQAAHAPVTLRASFADRTL 223

Qy 61 AWTSRWGOKKLISAPFLFYQRFPEVLLFOEGGLKALNGQGLGATSMNQALEQUREL 120  
Db 224 AWTSRWGOKKLISAPFLFYQRFPEVLLFOEGGLKALNGQGLGATSMNQALEQUREL 283

Qy 121 RISGSVOLYCVHS 133  
Db 284 RISGSVOLYCVHS 296

RESULT 10  
AAM48755  
ID AAM48755 standard; protein; 314 AA.  
XX  
AC AAM48755;  
XX  
DT 03-APR-2002 (first entry)  
XX  
DE Human adipose tissue specific polypeptide SEQ ID NO 2.  
XX  
KW Human; adipose tissue.  
XX  
OS Homo sapiens.  
XX  
PN JP2001309787-A.  
XX  
PD 06-NOV-2001.  
XX  
XX 02-MAY-2000; 2000JP-00133327.  
XX  
XX 02-MAY-2000; 2000JP-00133327.  
XX  
XX (SUMO ) SUMITOMO CHEM CO LTD.  
XX  
XX WPI; 2002-134851/18.  
XX  
XX N-PSDB; ABA96479.  
XX  
XX Identification of adipose tissue comprises detecting, in a test tissue  
XX sample, an expression product of a sequence found in adipose tissue.  
XX  
XX Claim 1; Page 12-13; 17pp; Japanese.  
XX  
XX The invention relates to identification of adipose tissue comprising that  
XX an expression product of a gene having a sequence encoding a sequence of  
XX 31 amino acids (AAM48755), given in the specification, is detected. The  
XX method is used for the detection of adipose tissue. The present sequence  
XX is that of an adipose specific sequence of the invention  
XX  
SQ Sequence 314 AA;

Query Match 98.4%; Score 671; DB 5; Length 314;  
Best Local Similarity 99.2%; Pred. No. 1.9e-75;  
Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSPRLVPCSHALPQGLSPGQVIIVRGLVLPQPKHPTVSLRDQAAHAPVTLRASFADRTL 60  
Db 182 MSPRLVPCSHALPQGLSPGQVIIVRGLVLPQPKHPTVSLRDQAAHAPVTLRASFADRTL 241

Qy 61 AWTSRWGOKKLISAPFLFYQRFPEVLLFOEGGLKALNGQGLGATSMNQALEQUREL 120  
Db 242 AWTSRWGOKKLISAPFLFYQRFPEVLLFOEGGLKALNGQGLGATSMNQALEQUREL 301

Qy 121 RISGSVOLYCVHS 133  
Db 302 RISGSVOLYCVHS 314

RESULT 11  
AAY70029  
ID AAY70029 standard; protein; 336 AA.  
XX  
AC AAY70029;  
XX  
DT 05-JUN-2000 (first entry)  
XX  
DE Human extracellular adhesive protein, EXADH1.  
XX  
KW Human; extracellular adhesive protein; EXADH1; cytostatic; antiarthritic;  
KW antiasthmatic; immunosuppressive; antiarteriosclerotic; diagnosis;  
KW treatment; prevention; cancer; leukaemia; melanoma; immune disorder;  
KW rheumatoid arthritis; asthma; atherosclerosis; developmental disorder;  
KW anaemia; epilepsy; Cushing's syndrome; cell growth disorder;  
KW cell differentiation; embryogenesis; morphogenesis; drug screening.  
XX  
OS Homo sapiens.  
XX  
XX  
XX Key Location/Qualifiers  
XX Modified-site 24  
XX Region /note= "Casein kinase II phosphorylation site"  
XX 48..174  
XX /note= "contains at least 46 out of 55 conserved CRD  
XX residues"  
XX 67..174  
XX /note= "This region is similar to galectin carbohydrate  
XX recognition domain (CRD)"  
XX Modified-site 102  
XX Modified-site /note= "Protein kinase C phosphorylation site"  
XX 132  
XX /note= "cAMP and cGMP-dependent protein kinase  
XX phosphorylation site"  
XX Modified-site 164  
XX /note= "Casein kinase II phosphorylation site"  
XX Modified-site 205  
XX /note= "Protein kinase C phosphorylation site"  
XX Modified-site 242  
XX /note= "Protein kinase C phosphorylation site"  
XX Modified-site 242  
XX /note= "Casein kinase II phosphorylation site"  
XX Modified-site 253  
XX /note= "Protein kinase C phosphorylation site"  
XX Modified-site 257  
XX /note= "Casein kinase II phosphorylation site"  
XX  
XX WO200009690-A1.  
XX  
XX 24-FEB-2000.  
XX  
XX 09-AUG-1999; 99WO-US017997.  
XX  
XX 10-AUG-1998; 98US-00131648.  
XX  
XX (INCY-) INCYTE PHARM INC.  
XX  
XX Hillman JJ, Yue H, Corley NC, Guegler KJ, Patterson C;  
XX WPI; 2000-224335/19.  
XX N-PSDB; AAZ50940.  
XX  
XX New human extracellular adhesive polypeptide and polynucleotide useful  
XX for diagnosis, prevention and treatment of cancer, immune disorders and  
XX PT

```

PT developmental disorders.
XX
PS Claim 1; Page 57-58; 67pp; English.
XX
CC The present sequence is an extracellular adhesive protein, EXADH1 which
CC was first identified in incyte clone 2635136 obtained from BONTNTO1 cDNA
CC library. The cDNA library was constructed using RNA isolated from tibial
CC periosteum removed from a 20 year old Caucasian male during a
CC hemipelvectomy with amputation above the knee. EXADH1 has chemical and
CC structural similarities with human prostate carcinoma tumour antigen 1
CC (PCTA-1). This sequence is expressed in libraries derived from
CC reproductive and haematopoietic tissues. The present sequence is useful
CC in the diagnosis, treatment and prevention of disorders associated with
CC abnormal expression of EXADH particularly cancers such as leukaemia and
CC melanoma, immune disorders such as rheumatoid arthritis, asthma and
CC atherosclerosis, and developmental disorders such as anaemia, epilepsy,
CC Cushing's syndrome and any disorder associated with cell growth,
CC differentiation, embryogenesis and morphogenesis. EXADH, its catalytic or
XX immunogenic fragments are also useful for drug screening
XX
SQ Sequence 336 AA;

Query Match 98.4%; Score 671; DB 3; Length 336;
Best Local Similarity 99.2%; Pred. No. 2.1e-75;
Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSPRLVPCSHALPQGLSPGVIIIVRGLVLOEPKHTVSLRDQAAHAPVTLRASPADRTL 60
DB 204 MSPRLVPCSHALPQGLSPGVIIIVRGLVLOEPKHTVSLRDQAAHAPVTLRASPADRTL 263
QY 61 AMISRWGQKLLISAPFLFYPQRFPEVLLLFQEGGLKALNGQGLGATSMNQALEQLREL 120
DB 264 AMISRWGQKLLISAPFLFYPQRFPEVLLLFQEGGLKALNGQGLGATSMNQALEQLREL 323
QY 121 RISGSVOLYCVHS 133
DB 324 RISGSVOLYCVHS 336

RESULT 12
AD113483
XX ID AD113483 standard; protein; 318 AA.
XX AC AD113483;
XX DT 22-APR-2004 (first entry)
XX DE Human galectin-8 protein sequence SeqID 4.
XX KW human; galectin-8; extracellular matrix; ECM; S-type lectin; cell growth;
XX metastatic potential; proliferative disorder; cytostatic; cancer; tumour.
XX OS Homo sapiens.
XX PN US2003220231-A1.
XX PD 27-NOV-2003.
XX PF 21-MAY-2002; 2002US-00151166.
XX PR 21-MAY-2002; 2002US-00151166.
XX PA (ZICK/) ZICK Y.
XX PA (LEVY/) LEVY Y.
XX PA (ARBE/) ARBEL-GOREN R.
XX PA (RONE/) RONE D.
XX PA (HADA/) HADARI Y R.
XX PI Zick Y, Levy Y, Arbel-Goren R, Ronen D, Hadari YR;
XX WPI; 2004-051718/05.
XX DR N-PSDB; AD113482.
XX

PT Use of galectin-8 and galectin-8-like proteins or anti-galectin-8
PT antibodies, for modulating cell adhesion or treating a pathological
PT disorder related to abnormal cell growth, e.g. lung, liver, brain,
PT bladder or prostate cancer.
XX
PS Claim 1; SEQ ID NO 4; 28pp; English.
XX
CC This invention relates to novel compositions that enhance or inhibit the
CC interactions of galectin-8 (and galectin-8-like proteins) with
CC extracellular matrix (ECM) proteins or cell surface receptors.
CC Specifically, it refers to a method for modulating cell adhesion using
CC these proteins, antisense oligos or derived anti-galectin-8 antibodies
CC thereof. The present invention describes human galectin-8, an S-type
CC lectin, as a secreted protein that is involved in the regulation of cell
CC growth and development, including metastatic potential, such that it
CC plays a role in modulating cell-matrix interactions in pathological
CC processes including tumour development and various proliferative
CC disorders. Accordingly, compositions of this invention exhibit cytostatic
CC activities and can be used to treat abnormal cell growth i.e. cancer and
CC tumours both in vivo or ex vivo. This polypeptide is the human galectin-8
XX protein sequence of the invention.
XX
SQ Sequence 318 AA;

Query Match 18.3%; Score 124.5; DB 8; Length 318;
Best Local Similarity 25.5%; Pred. No. 1.4e-06;
Matches 35; Conservative 24; Mismatches 73; Indels 5; Gaps 1;

QY 2 SPRLEVPCHALPQGLSPGVIIIVRGLVLOEPKHTVSL-----RDQAAHAPVTLRASPA 56
DB 181 TPQLRLPFAALTPMGPGRTVVVKGEVMAKSPNVLLAGKSKDIALHLNPLNIKAP 240
QY 57 DRTLANISRWGQKLLISAPFLFYPQRFPEVLLLFQEGGLKALNGQGLGATSMNQALEQL 116
DB 241 VRNSPLQESWGSEERNITSPFSPGMYPHMIITCDVREFKVAVNGVHSLEYKRPKELSS 300
QY 117 LRELISGSVOLYCVHS 133
DB 301 IDTLINGDIHLLEVR 317

RESULT 13
AA79505
XX ID AA79505 standard; protein; 171 AA.
XX AC AA79505;
XX DT 01-AUG-2000 (first entry)
XX DE Human carbohydrate-associated protein CRBAP-1.
XX KW CRBAP-1; carbohydrate-associated protein 1; human; autoimmune disorder;
XX inflammation; gastrointestinal disorder; infection;
XX reproductive disorder; neurological disorder; eye disorder;
XX cell proliferation; cancer; diagnosis; therapy.
XX OS Homo sapiens.
XX PH Key Location/Qualifiers
XX Modified-site 4 /note= "O-phosphorylated"
XX Modified-site 20 /note= "N-glycosylated"
XX Peptide 58..165 /note= "signature peptide"
XX Modified-site 72 /note= "O-phosphorylated"
XX Modified-site 92 /note= "O-phosphorylated"
XX Modified-site 155 /note= "O-phosphorylated"
XX Modified-site 159 /note= "O-phosphorylated"
XX Modified-site /note= "O-phosphorylated"

```

XX WO200018922-A2.  
 PN 06-APR-2000.  
 PD 29-SEP-1999; 99WO-US022585.  
 XX 01-OCT-1998; 98US-00164785.  
 PR 01-OCT-1998; 98US-0155267P.  
 PR 06-OCT-1998; 98US-00167179.  
 PR 06-OCT-1998; 98US-0155266P.  
 PR 13-NOV-1998; 98US-00191838.  
 PR 13-NOV-1998; 98US-0155227P.  
 PR 03-DEC-1998; 98US-00205656.  
 XX (INCYTE PHARM INC.  
 PA Au-Young J, Lal P, Bandman O, Reddy R, Baughn MR, Yue H;  
 PI Hillman JL;  
 XX WPI: 2000-317516/27.  
 DR N-PSDB; AAZ94941, AAZ94948.  
 XX Novel carbohydrate-associated proteins used for the prevention and  
 PT treatment of autoimmune/inflammatory disorders of e.g. the  
 PT gastrointestinal and reproductive systems.  
 XX Claim 1; Page 82-83; 104pp; English.  
 XX The present sequence is that of a novel human carbohydrate-associated  
 CC protein, termed CRBAP-1, as deduced from cDNA (see AAZ94941) obtained  
 CC from prostate tumour cDNA library PROSTUT01. CRBAP-1 resembles vertebrate  
 CC galactoside-binding lectins in the C58-D165 region, showing 28% identity  
 CC to human galectin-8. CRBAP-1 expression was identified in libraries  
 CC associated with cancer and cell proliferation, inflammation and immune  
 CC response, reproductive tissues and nervous tissues. The invention  
 CC provides CRBAP-1 to -7 polynucleotides (see AAZ94941-48) and polypeptides  
 CC (see AAY79505-11), as well as expression vectors, host cells, antibodies,  
 CC agonists and antagonists. These are used in the diagnosis, treatment or  
 CC prevention of disorders associated with CRBAP expression, especially  
 CC autoimmune or inflammatory disorders, gastrointestinal disorders,  
 CC infectious disorders, reproductive disorders, neurological disorders, eye  
 CC disorders and cell proliferative disorders, including cancer  
 XX Sequence 171 AA;  
 SQ  
 Query Match 17.7%; Score 120.5; DB 3; Length 171;  
 Best Local Similarity 29.4%; Pred. No. 2e-06;  
 Matches 40; Conservative 27; Mismatches 58; Indels 11; Gaps 5;  
 Qy 3 PRLEVP-CSHALPQGLSPGQVIVRGVLQEPKHPVTSL-----RDQAAHAPVTLRASFA 56  
 Db 34 PRLIVPFCGH-IKGMRPGKVLVNGIVDLNPESFAISLTGDSDEPPADVAIELKAVFT 92  
 Qy 57 DRTL---AWIS-RWGQKKLISAPFLYPQRFVEVLLLFQEGGLKALNGOGLGATSMNQ 112  
 Db 93 DRLLRNSCISGERGEQSAIPYFPFIPDQPRVEILCEHPRFRVFDVGHQLDFYHRIQ 152  
 Qy 113 ALEQLRELIRISGSVOL 128  
 Db 153 TLSAIDTIKINGDLQI 168  
 RESULT 14  
 ID AAB70086  
 XX AAB70086 standard; protein; 196 AA.  
 AC AAB70086;  
 XX 14-MAY-2001 (first entry)  
 DT Human secreted protein-related polypeptide #7.  
 DE  
 XX

Human; secreted protein; immunomodulatory; antisclerotic; dermatological;  
 immunosuppressive; antiinflammatory; anti-HIV; immunostimulant;  
 cytostatic; cardiac; vascular; anti-angiogenic; ophthalmological;  
 neuroprotectant; nootropic; anticonvulsant; vaccine; antialzheimers;  
 antiparkinsonian; antimicrobial; vulnery; gene therapy;  
 immune disorder; hyperproliferative; cardiovascular; angiogenic;  
 neurological; infection.  
 Homo sapiens.  
 WO200112776-A2.  
 22-FEB-2001.  
 15-AUG-2000; 2000WO-US022350.  
 16-AUG-1999; 99US-0148759P.  
 (HUMA-) HUMAN GENOME SCI INC.  
 Shi Y, Young PE, Ebner R, Soppet DR, Ruben SM;  
 WPI: 2001-244245/25.  
 Nucleic acids encoding 18 human secreted polypeptides, useful for  
 preventing, diagnosing and/or treating e.g. cancers, Parkinson's disease  
 and diabetic retinopathy.  
 Disclosure; Page 378; 380pp; English.  
 The present sequence is given in a specification relating to nucleic acid  
 molecules encoding 18 novel human secreted proteins. The nucleic acids  
 and proteins may be used in the prevention, diagnosis and treatment of  
 diseases including immune disorders (e.g. multiple sclerosis, systemic  
 lupus erythematosus and human immunodeficiency virus (HIV) infections),  
 hyperproliferative disorders (e.g. cancers and Gaucher's disease),  
 cardiovascular diseases (e.g. Schmitz syndrome, Chaga's cardiomyopathy  
 and coronary arteriosclerosis), angiogenic disorders (e.g. corneal graft  
 neovascularisation and diabetic retinopathy), neurological disorders  
 (e.g. Huntington's chorea, Alzheimer's disease and Parkinson's disease),  
 infectious diseases and/or for promoting wound healing, regeneration  
 and/or chemotaxis. The nucleic acid molecules may also be used as DNA  
 probes in diagnostic assays to detect and quantitate the presence of  
 similar nucleic acid sequences in samples. The polypeptides may also be  
 used as antigens in the production of antibodies and in assays to  
 identify modulators of protein expression and activity  
 Sequence 196 AA;  
 SQ  
 Query Match 17.7%; Score 120.5; DB 4; Length 196;  
 Best Local Similarity 29.4%; Pred. No. 2.4e-06;  
 Matches 40; Conservative 27; Mismatches 58; Indels 11; Gaps 5;  
 Qy 3 PRLEVP-CSHALPQGLSPGQVIVRGVLQEPKHPVTSL-----RDQAAHAPVTLRASFA 56  
 Db 58 PRLIVPFCGH-IKGMRPGKVLVNGIVDLNPESFAISLTGDSDEPPADVAIELKAVFT 116  
 Qy 57 DRTL---AWIS-RWGQKKLISAPFLYPQRFVEVLLLFQEGGLKALNGOGLGATSMNQ 112  
 Db 117 DRLLRNSCISGERGEQSAIPYFPFIPDQPRVEILCEHPRFRVFDVGHQLDFYHRIQ 176  
 Qy 113 ALEQLRELIRISGSVOL 128  
 Db 177 TLSAIDTIKINGDLQI 192  
 RESULT 15  
 ID ADE57164  
 XX ADE57164 standard; protein; 144 AA.  
 AC ADE57164;  
 XX 29-JAN-2004 (first entry)  
 DT



XX DE Rat Protein AAA65445, SEQ ID NO 3025.  
XX KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
XX KW Chronic constriction injury; CCI; spared nerve injury; SNI; Chung.  
XX OS Rattus norvegicus.  
XX PN WO2003016475-A2.  
XX PD 27-FEB-2003.  
XX PF 14-AUG-2002; 2002WO-US025765.  
XX PR 14-AUG-2001; 2001US-0312147P.  
XX PR 01-NOV-2001; 2001US-0346382P.  
XX PR 26-NOV-2001; 2001US-0333347P.  
XX PA (GEHO ) GEN HOSPITAL CORP.  
XX PA (FARB ) BAYER AG.  
XX PI Woolf C, D'urso D, Befort K, Costigan M;  
XX PI WPI; 2003-268312/26.  
XX DR GENBANK; AAA65445.  
XX PT New composition comprising two or more isolated polypeptides, useful for  
XX PT preparing a medicament for treating pain in an animal.  
XX PS Claim 1; Page; 1017pp; English.  
XX CC The invention discloses a composition comprising two or more isolated rat  
XX CC or human polynucleotides or a polynucleotide which represents a fragment,  
XX CC derivative or allelic variation of the nucleic acid sequence. Also  
XX CC claimed are a vector comprising the novel polynucleotide, a host cell  
XX CC comprising the vector, a method for identifying a nucleotide sequence  
XX CC which is differentially regulated in an animal subjected to pain and a  
XX CC kit to perform the method, an array, a method for identifying an agent  
XX CC that increases or decreases the expression of the polynucleotide sequence  
XX CC that is differentially expressed in neuronal tissue of a first animal  
XX CC subjected to pain, a method for identifying a compound which regulates  
XX CC the expression of a polynucleotide sequence which is differentially  
XX CC expressed in an animal subjected to pain, a method for identifying a  
XX CC compound that regulates the activity of one or more of the  
XX CC polynucleotides, a method for producing a pharmaceutical composition, a  
XX CC method for identifying a compound or small molecule that regulates the  
XX CC activity in an animal of one or more of the polypeptides given in the  
XX CC specification, a method for identifying a compound useful in treating  
XX CC pain and a pharmaceutical composition comprising the one or more  
XX CC polypeptides or their antibodies. The polynucleotide or the compound that  
XX CC modulates its activity is useful for preparing a medicament for treating  
XX CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
XX CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
XX CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
XX CC the specification) which is differentially expressed during pain. Note:  
XX CC The sequence data for this patent did not form part of the printed  
XX CC specification, but was obtained in electronic form directly from WIPO at  
XX CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX CC Sequence 144 AA;

Query Match 17.4%; Score 119; DB 7; Length 144;  
Best Local Similarity 26.8%; Pred. No. 2.4e-06;  
Matches 35; Conservative 26; Mismatches 65; Indels 4; Gaps 2;

QY 3 PRLEVPCSHALPQGLSPGVIIIVRGLVLQEPKHTVSLR---DQAAHAPVTLRASFADRT 59  
DB 10 PNLAVFFTSIPNGLYPSKISVIGSVLSDAKRFQINLRCGGDIAFHLNPRFDENAVVRN 69  
QY 60 LAWISRWG-QKKLISAPFLFYPPQRFVEVLLFPEGGLKALNCGQIGLGTSMNQALEQLR 118  
DB 70 TQINNNGPERSLPGSMPPFSRQRFQSVWVLCBGHCFCFKAVDQGHICYSRLMLNLPDIN 129

QY 119 ELRISGSVQL 128  
DB 130 TLEVAGDIQL 139

RESULT 16  
ADE57162  
ID ADE57162 standard; protein; 144 AA.  
XX ADE57162;  
XX 29-JAN-2004 (first entry)  
XX Rat Protein AAA65445, SEQ ID NO 3023.  
DE Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
KW Chronic constriction injury; CCI; spared nerve injury; SNI; Chung.  
XX Rattus norvegicus.  
OS WO2003016475-A2.  
XX 27-FEB-2003.  
XX 14-AUG-2002; 2002WO-US025765.  
XX 14-AUG-2001; 2001US-0312147P.  
XX 01-NOV-2001; 2001US-0346382P.  
XX 26-NOV-2001; 2001US-0333347P.  
XX (GEHO ) GEN HOSPITAL CORP.  
XX PA (FARB ) BAYER AG.  
XX Woolf C, D'urso D, Befort K, Costigan M;  
XX WPI; 2003-268312/26.  
XX DR GENBANK; AAA65445.  
XX PT New composition comprising two or more isolated polypeptides, useful for  
XX PT preparing a medicament for treating pain in an animal.  
XX PS Claim 1; Page; 1017pp; English.  
XX CC The invention discloses a composition comprising two or more isolated rat  
XX CC or human polynucleotides or a polynucleotide which represents a fragment,  
XX CC derivative or allelic variation of the nucleic acid sequence. Also  
XX CC claimed are a vector comprising the novel polynucleotide, a host cell  
XX CC comprising the vector, a method for identifying a nucleotide sequence  
XX CC which is differentially regulated in an animal subjected to pain and a  
XX CC kit to perform the method, an array, a method for identifying an agent  
XX CC that increases or decreases the expression of the polynucleotide sequence  
XX CC that is differentially expressed in neuronal tissue of a first animal  
XX CC subjected to pain, a method for identifying a compound which regulates  
XX CC the expression of a polynucleotide sequence which is differentially  
XX CC expressed in an animal subjected to pain, a method for identifying a  
XX CC compound that regulates the activity of one or more of the  
XX CC polynucleotides, a method for producing a pharmaceutical composition, a  
XX CC method for identifying a compound or small molecule that regulates the  
XX CC activity in an animal of one or more of the polypeptides given in the  
XX CC specification, a method for identifying a compound useful in treating  
XX CC pain and a pharmaceutical composition comprising the one or more  
XX CC polypeptides or their antibodies. The polynucleotide or the compound that  
XX CC modulates its activity is useful for preparing a medicament for treating  
XX CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
XX CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
XX CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
XX CC the specification) which is differentially expressed during pain. Note:  
XX CC The sequence data for this patent did not form part of the printed  
XX CC specification, but was obtained in electronic form directly from WIPO at  
XX CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX CC Sequence 144 AA;



	Matches	35;	Conservative	26;	Mismatches	65;	Indels	4;	Gaps	2;
Qy	3	PRLVPCSHALPGCLSPGGQVIIVRGLVLQBPKHPTVSRL---	DOAAHAPVTTLRASFPADRT	59						
			:::	:::	:::	:::	:	:	:	:
			:::	:::	:::	:::	:	:	:	:
Dd	11	PNLAVPFTTSIPNGLYPSKSIIVSVGLSDAKRFQINRCGGDIAPHLNPRFDENAVVRN	70							
			:::	:::	:::	:::	:	:	:	:
Qy	60	LAWISRWG-OKKLISAPFLFYPPORPFTEVLLLFQEGGLKLALNGGGLGATSMNQQALEQLR	118							
		: :::	:::	:::	:::	:::	:	:	:	:
Dd	71	TQINNWSGPBESRLPGSMPFSRQGQFVSWILTCGHCFKVAVDGQHICEYSHRLMLNPDLIN	130							
			:::	:::	:::	:::	:	:	:	:
Qy	119	ELRISGSVOL	128							
			:::	:::	:::	:::	:	:	:	:
Dd	131	TLEVAGDIQL	140							
			:::	:::	:::	:::	:	:	:	:

RESULT 18	
ADG40248	
ID	ADG40248 standard; protein; 145 AA.
XX	
AC	ADG40248;
XX	
DT	26-FEB-2004 (first entry)
XX	
DE	Rat galectin 5.
XX	
KW	Galectin 11; cytostatic; immunosuppressive; antiinflammatory;
KW	antiallergic; antiallergic; gene therapy; cell growth disorder; cancer;
KW	autoimmune disease; inflammatory disease; asthma; allergic disease; rat.
XX	
OS	Rattus sp.
XX	
FN	US2003208044-A1.
XX	
PD	06-NOV-2003.
XX	
PF	06-JUN-2003; 2003US-00455366.
XX	
PR	21-JAN-1997; 97US-0034204P.
PR	21-JAN-1997; 97US-0034205P.
PR	21-JAN-1998; 98US-00010146.
PR	06-JUN-1998; 98US-00109864.
PR	21-APR-1999; 99US-0130390P.
PR	10-DEC-1999; 99US-0169932P.
PR	21-APR-2000; 2000US-00557170.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
PA	(JOL-) LA JOLLA INST ALLERGY & IMMUNOLOGY.
XX	
PI	Ni J, Gentz RL, Rosen CA, Liu F;
XX	
DR	WPI; 2003-864797/80.
XX	
FT	New galectin 11 polypeptides, useful for treating a cell growth disorder,
PT	e.g. cancer, autoimmune diseases, inflammatory diseases, asthma, or
PT	allergic diseases.
XX	
PS	Disclosure; SEQ ID NO 3; 123pp; English.
XX	
CC	The invention relates to an isolated polynucleotide comprising a sequence
CC	that is at least 95% identical to a sequence encoding Human Galectin 11,
CC	galectin 11alpha, galectin 11beta (or their defined fragments). Also
CC	included are a method of making a recombinant vector by inserting the
CC	isolated polynucleotide into a vector, a recombinant vector comprising the
CC	polynucleotide, a genetically engineered host cell comprising the
CC	polynucleotide, a method of producing a galectin 11 polypeptide (by
CC	culturing the genetically engineered host cell under conditions suitable
CC	to produce the polypeptide, and recovering the polypeptide), an isolated
CC	galectin 11 polypeptide comprising any of the amino acid sequences cited
CC	above, a pharmaceutical composition comprising the polypeptide and a
CC	carrier, an isolated antibody that binds specifically to the polypeptide,
CC	a method of detecting a galectin 11 polypeptide in a sample (by
CC	contacting the sample with the antibody, and detecting the presence of
CC	the antibody bound to the polypeptide), a method of treating a cell

CC growth disorder in a mammal by administering the polypeptide to the  
CC mammal, and a method of regulating cell growth or differentiation in a  
CC mammal by administering a galectin II polypeptide or polynucleotide to  
CC the mammal to suppress cell growth or differentiation. The  
CC polynucleotide, polypeptide and composition are useful for treating a  
CC cell growth disorder, e.g. cancer, autoimmune diseases, inflammatory  
CC diseases, asthma, or allergic diseases (many examples cited in the  
CC specification). The present sequence represents a homologue of a galectin  
CC II protein.

XX  
SQ Sequence 145 AA;

Query Match 17.4%; Score 119; DB 7; Length 145;  
Best Local Similarity 26.9%; Pred. No. 2.4e-06;  
Matches 35; Conservative 26; Mismatches 65; Indels 4; Gaps 2;

QY 3 PRLEVPCHALPQGLSPGVIIIRGLVLPQPKHPTVSLR---DQAAHAPVTLRASFADRT 59  
DB 11 PNLAVPFTTSIPNGLYPSKSIIVISGVVLSDAKRFQINLRCGGDIAFHLNPRFDENAVVRN 70  
QY 60 LAWISRWG-OKKLISAPFLFYPPQRFVLLFOEGGLKALNGQGLGATSMNQALEQLR 118  
DB 71 TQINNSWGPERSLPGSMPPSRQRFVSWILCEGHCFKVAVDGQHICEYSHRLMNLPDIN 130  
QY 119 ELRISGSVOL 128  
DB 131 TLEVAGDIQL 140

RESULT 19  
ADN04944  
ID ADN04944 standard; protein; 172 AA.

XX AC ADN04944;  
XX DT 01-JUL-2004 (first entry)  
XX DE Antipsoriatic protein sequence #652.  
XX KW antipsoriatic; gene therapy; psoriasis; diagnosis.  
XX OS Homo sapiens.  
XX PN WO2004028479-A2.  
XX PD 08-APR-2004.  
XX PF 25-SEP-2003; 2003WO-US030907.  
XX PR 25-SEP-2002; 2002US-0414006P.  
XX PA (GETH ) GENENTECH INC.  
XX PI Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WT;  
XX PI Wu TD;  
XX WPI; 2004-305105/28.  
XX DR N-PSDB; ADN04943.

XX PT New PRO nucleic acid or polypeptide, useful for preparing a  
XX PT pharmaceutical composition for diagnosing or treating psoriasis in a  
XX PT mammal.  
XX PS Claim 9; SEQ ID NO 1338; 3069pp; English.  
XX CC The invention relates to novel polynucleotide and polypeptides for  
XX CC treating psoriasis or a sequence having at least 80% identity to the  
XX CC above sequences. The nucleic acid is useful for preparing a composition  
XX CC for diagnosing or treating psoriasis in a mammal. This sequence  
XX CC corresponds to one of the polypeptides of the invention.  
XX SQ Sequence 172 AA;

Query Match 17.4%; Score 118.5; DB 8; Length 172;  
Best Local Similarity 29.4%; Pred. No. 3.6e-06;  
Matches 40; Conservative 27; Mismatches 58; Indels 11; Gaps 5;

QY 3 PRLEVP-CHALPQGLSPGVIIIRGLVLPQPKHPTVSL-----RQAAHAPVTLRASF 56  
DB 34 PRLIVPFCGH-IKGGMWPGKVKVIMGIVDLNPESFAISLTGDSDEPPADVAIELKAVFT 92  
QY 57 DRTL---AMIS-RWGOKKLISAPFLFYPPQRFVLLFOEGGLKALNGQGLGATSMNQ 112  
DB 93 DRQLNNSCISGERGEEQSAIPYFPIDQPPRVEILCEYPRFRVFDGHLQFDFYHRIQ 152  
QY 113 ALEQLRELRISSVOL 128  
DB 153 TLSAIDTIKINGDLQI 168

RESULT 20  
ABR69609

ID ABR69609 standard; protein; 158 AA.

XX AC ABR69609;  
XX DT 12-AUG-2003 (first entry)  
XX DE Human CGDD-9 protein.  
XX KW Human; cytostatic; antiarteriosclerotic; anticonvulsant; nootropic;  
KW neuroprotective; cerebroprotective; anti-HIV; antiallergic;  
KW antiinflammatory; gynaecological; cancer; atherosclerosis; epilepsy;  
KW Huntington's disease; stroke; AIDS; allergy; placenta; reproductive;  
KW CGDD; cell growth; cell differentiation; cell death.  
XX OS Homo sapiens.  
XX PN WO2003027263-A2.  
XX PD 03-APR-2003.  
XX PF 26-SEP-2002; 2002WO-US031095.  
XX PR 28-SEP-2001; 2001US-0326389P.  
XX PR 05-OCT-2001; 2001US-0327380P.  
XX PR 05-OCT-2001; 2001US-0328186P.  
XX PR 12-OCT-2001; 2001US-0329690P.  
XX PR 26-OCT-2001; 2001US-0345384P.  
XX PR 26-OCT-2001; 2001US-0348165P.  
XX PR 02-NOV-2001; 2001US-0350219P.  
XX PR 09-NOV-2001; 2001US-034518P.  
XX PR 09-NOV-2001; 2001US-0345143P.  
XX PR 16-NOV-2001; 2001US-0332375P.  
XX PR 03-DEC-2001; 2001US-0336908P.  
XX PR 07-DEC-2001; 2001US-0340747P.  
XX PA (INCY-) INCYTE GENOMICS INC.

XX PI Azimzai Y, Baughn MR, Becha SD, Borowsky ML, Chawla NK;  
XX PI Elliott JS, Emerling BM, Gandhi AR, Gietzen KJ, Gorvad AE;  
XX PI Griffin VA, Hafalia AJA, Ison CH, Kable AE, Kalafus DP;  
XX PI Lehr-Mason TM, Lu DAM, Marquis JP, Nguyen DB, Ramkumar J;  
XX PI Richardson PW, Sapperstein SK, Swarnakar A, Tang YT, Tran UK;  
XX PI Warren BA, Xu Y, Yao MG, Yue H, Yue H;  
XX WPI; 2003-421159/39.  
XX DR N-PSDB; ACC90586.  
XX PT New human proteins associated with cell growth, differentiation, and  
XX PT death (CGDD), useful for diagnosing, treating and preventing diseases or  
XX PT conditions associated with the aberrant CGDD expression e.g. cancer,  
XX PT AIDS, or epilepsy.  
XX PS Claim 1; Page 257; 350pp; English.

CC The invention relates to an isolated polypeptide associated with cell growth, differentiation and death (CGDD). Also disclosed are the polynucleotides encoding the polypeptides. The polypeptides and polynucleotides are useful in diagnosing, treating and preventing diseases or conditions associated with the decreased expression or over expression of CGDD. Such diseases include cell proliferative (e.g. cancer, atherosclerosis), neurological (e.g. epilepsy, Huntington's disease, stroke), immune/inflammatory (e.g. AIDS, allergies) and reproductive disorders, or disorders of the placenta. They are also useful in assessing the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of CGDD. The CGDD or its fragments are useful in screening compounds for effectiveness as an agonist or antagonist of the polypeptides, or in altering the expression of the target polynucleotide and compounds that specifically bind to or modulate the activity of the polypeptide. Microarrays consisting of polynucleotides of the invention are useful in monitoring or measuring protein-protein interactions, drug-target interactions, and gene expression profiles. Sequences given in records AAR69601-AAR69657 represent CGDD polypeptides of the invention

XX SQ Sequence 158 AA;

Query Match 16.1%; Score 109.5; DB 6; Length 158;  
Best Local Similarity 25.5%; Pred. No. 4.3e-05;  
Matches 35; Conservative 23; Mismatches 74; Indels 5; Gaps 1;

QY 2 SPRLEVPCHALPQGLSPGQVIVRGVLQEPKHTVSL-----RDQAAHAPVTLRASFA 56  
DB 21 TPQLSLPFAARLNTPMGPGTGVVVKGEVNAKSNFVNDLLAGKSKDIALHNLPRINIKAF 80

QY 57 DRTLAWISRWGQKLLISAPFLFPQRFVLLLFQEGGLKALNGQGLGATSMNQALEQ 116  
DB 81 VRNSFLOESNGEERNITSPFSPGMYFEMIYCDVREFKAVNGVHSLVKHFKELSS 140

QY 117 LRELIRSGSVOLYCVHS 133  
DB 141 IDTLEINGDIHLLVRS 157

RESULT 21  
AD021123  
ID AD021123 standard; protein; 258 AA.

XX AC AD021123;  
XX DT 12-AUG-2004 (first entry)  
XX DE Human carbohydrate-associated protein (CHOP) #19.  
XX KW human; carbohydrate-associated protein; CHOP; neurodegenerative disorder;  
XX KW Parkinson's disease; Alzheimer's disease; muscular disorder;  
XX KW myotonic dystrophy; catatonias; endocrine disorder; diabetes;  
XX KW Grave's disease; cancer; leukemia; immunological disorder; scleroderma;  
XX KW systemic lupus erythematosus; allergy; gastrointestinal disorder;  
XX KW Crohn's disease; renal disorder; Goodpasture's syndrome; infection;  
XX KW cardiovascular disorder; atherosclerosis; hepatic disease; cirrhosis.

XX OS Homo sapiens.  
XX PN WO2004041166-A2.  
XX PD 27-MAY-2004.  
XX PF 10-NOV-2003; 2003WO-US035947.  
XX PR 12-NOV-2002; 2002US-0425423P.  
XX PR 21-JAN-2003; 2003US-0441847P.  
XX PR 10-MAR-2003; 2003US-0453882P.  
XX PR 20-MAR-2003; 2003US-0456645P.  
XX PR 16-APR-2003; 2003US-0463676P.  
XX PA (INCY-) INCYTE CORP.

PI Chawla NK, Tang YT, Griffin JA, Yang YG, Ramkumar J, Khare R, Richardson TW, Becha SD, Tran UK, Kable AE, Swarnakar A, Warren BA; Elliott VS, Marquis JP, Hafalia AJA;  
XX WPI: 2004-420308/39.  
DR N-PSDB; ADO21143.  
XX New CHOP polypeptides, useful for diagnosing, preventing, and treating disorders associated with abnormal expression or activity of CHOP, e.g. neuromuscular, immunological, cardiovascular disorders, cancer and/or infections.

PS Claim 1; SEQ ID NO 19; 181pp; English.

XX The invention comprises the amino acid and coding sequences of human carbohydrate-associated proteins (CHOP). The DNA and protein sequences of the invention are useful for diagnosing, preventing and treating disorders associated with abnormal expression or activity of CHOP, such as: neurodegenerative disorders (e.g. Parkinson's disease, Alzheimer's disease), muscular disorders (e.g. myotonic dystrophy, catatonias), endocrine disorders (e.g. diabetes, Grave's disease), cancers (e.g. leukemia), immunological disorders (e.g. scleroderma, systemic lupus erythematosus, allergies), gastrointestinal disorders (e.g. Crohn's disease), renal disorders (e.g. Goodpasture's syndrome), infections (e.g. viral, bacterial, fungal, parasitic, protozoal, helminthic), cardiovascular disorders (e.g. atherosclerosis), or hepatic diseases (e.g. cirrhosis). The present amino acid sequence represents a human CHOP protein of the invention.

XX SQ Sequence 258 AA;

Query Match 16.1%; Score 109.5; DB 8; Length 258;  
Best Local Similarity 25.5%; Pred. No. 8.4e-05;  
Matches 35; Conservative 23; Mismatches 74; Indels 5; Gaps 1;

QY 2 SPRLEVPCHALPQGLSPGQVIVRGVLQEPKHTVSL-----RDQAAHAPVTLRASFA 56  
DB 121 TPQLSLPFAARLNTPMGPGTGVVVKGEVNAKSNFVNDLLAGKSKDIALHNLPRINIKAF 180

QY 57 DRTLAWISRWGQKLLISAPFLFPQRFVLLLFQEGGLKALNGQGLGATSMNQALEQ 116  
DB 181 VRNSFLOESNGEERNITSPFSPGMYFEMIYCDVREFKAVNGVHSLVKHFKELSS 240

QY 117 LRELIRSGSVOLYCVHS 133  
DB 241 IDTLEINGDIHLLVRS 257

RESULT 22  
ABM83347  
ID ABM83347 standard; protein; 258 AA.  
XX AC ABM83347;  
XX DT 18-NOV-2004 (first entry)  
XX DE Human diagnostic and therapeutic pprotein SEQ ID NO:3596.  
XX KW gene therapy; human diagnostic and therapeutic polynucleotide; dithp.  
XX OS Homo sapiens.  
XX PN WO2004023973-A2.  
XX PD 25-MAR-2004.  
XX PR 12-SEP-2003; 2003WO-US028227.  
XX PR 12-SEP-2002; 2002US-0410259P.  
XX PR 12-SEP-2002; 2002US-0410260P.  
XX PA (INCY-) INCYTE CORP.

PI Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;  
PI Harthshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;  
PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;  
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;  
PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;  
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirtzen D;  
PI Xu Y, Kwong M, Pollicky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;  
PI Patury S, Shi X, Suarez CJ;  
XX WPI: 2004-329368/30.  
DR N-PSDB; ACM41999.  
DR XX  
XX New diagnostic and therapeutic polynucleotides and polypeptides, useful  
PT in diagnosing a condition, disease or disorder associated with human  
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or  
PT in gene mapping.  
XX  
PS Claim 27; Page; 190pp; English.  
XX  
CC The invention relates to novel diagnostic and therapeutic polynucleotides  
CC selected from one of the 2722 sequences defined in the specification. A  
CC polynucleotide of the invention may have a use in gene therapy. The human  
CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be  
CC used to diagnose a particular condition, disease or disorder associated  
CC with human molecules, e.g. cell proliferative disorders,  
CC autoimmune/inflammatory disorder, developmental disorder, endocrine  
CC disorder, neurological disorders, gastrointestinal disorders, or  
CC infections caused by virus, bacteria, fungi or parasite. The dithp  
CC molecules may also be used in genetic mapping, in identifying individuals  
CC from minute biological samples, in detecting single nucleotide  
CC polymorphisms, as molecular weight markers, and for somatic or germline  
CC gene therapy. The present sequence represents a dithp protein of the  
CC invention. Note: The sequence data for this patent is not represented in  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at [www.wipo.int/pct/en/sequences/listing.htm](http://www.wipo.int/pct/en/sequences/listing.htm)  
XX  
SQ Sequence 258 AA;  
Query Match 16.1%; Score 109.5; DB 8; Length 258;  
Best Local Similarity 25.5%; Pred. No. 8.4e-05;  
Matches 35; Conservative 23; Mismatches 74; Indels 5; Gaps 1;  
QY 2 SPRLEVPCHALPQGLSPGQVIVRGLVQEPKHTVSL-----RDQAAHAPVTLRASFA 56  
Db 121 TPQLRLPFAARLNTPMGPGRVTVVKEVNAVNAKSFNVDDLAKSKDIALHLNPLNIKAF 180  
QY 57 DRTLAWISRWGQKKLISAPFLFYPPQFFVLLFOEGGLKALNGQGLGATSNMQOALEQ 116  
Db 181 VRNSFLQESWGEEERNITSPFSPGMYFEMIIYCDVREFKAVNGVHSLEYKHFRELSS 240  
QY 117 LRELIRSGSVQLYCVHS 133  
Db 241 IDTLEINGDIHLLEVR 257  
RESULT 23  
ABM83348  
ID ABM83348 standard; protein; 258 AA.  
XX AC ABM83348;  
XX DT 18-NOV-2004 (first entry)  
XX DE Human diagnostic and therapeutic pprotein SEQ ID NO:3597.  
XX KW gene therapy; human diagnostic and therapeutic polynucleotide; dithp.  
XX OS Homo sapiens.  
XX XX WO2004023973-A2.  
XX PD 25-MAR-2004.  
XX

PF 12-SEP-2003; 2003WO-US028227.  
XX  
PR 12-SEP-2002; 2002US-0410259P.  
PR 12-SEP-2002; 2002US-0410260P.  
XX (INCY-) INCYTE CORP.  
PA Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;  
XX Harthshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;  
PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;  
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;  
PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;  
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirtzen D;  
PI Xu Y, Kwong M, Pollicky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;  
PI Patury S, Shi X, Suarez CJ;  
XX WPI: 2004-329368/30.  
DR N-PSDB; ACM42000.  
DR XX  
XX New diagnostic and therapeutic polynucleotides and polypeptides, useful  
PT in diagnosing a condition, disease or disorder associated with human  
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or  
PT in gene mapping.  
XX  
PS Claim 27; Page; 190pp; English.  
XX  
CC The invention relates to novel diagnostic and therapeutic polynucleotides  
CC selected from one of the 2722 sequences defined in the specification. A  
CC polynucleotide of the invention may have a use in gene therapy. The human  
CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be  
CC used to diagnose a particular condition, disease or disorder associated  
CC with human molecules, e.g. cell proliferative disorders,  
CC autoimmune/inflammatory disorder, developmental disorder, endocrine  
CC disorder, neurological disorders, gastrointestinal disorders, or  
CC infections caused by virus, bacteria, fungi or parasite. The dithp  
CC molecules may also be used in genetic mapping, in identifying individuals  
CC from minute biological samples, in detecting single nucleotide  
CC polymorphisms, as molecular weight markers, and for somatic or germline  
CC gene therapy. The present sequence represents a dithp protein of the  
CC invention. Note: The sequence data for this patent is not represented in  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at [www.wipo.int/pct/en/sequences/listing.htm](http://www.wipo.int/pct/en/sequences/listing.htm)  
XX  
SQ Sequence 258 AA;  
Query Match 16.1%; Score 109.5; DB 8; Length 258;  
Best Local Similarity 25.5%; Pred. No. 8.4e-05;  
Matches 35; Conservative 23; Mismatches 74; Indels 5; Gaps 1;  
QY 2 SPRLEVPCHALPQGLSPGQVIVRGLVQEPKHTVSL-----RDQAAHAPVTLRASFA 56  
Db 121 TPQLRLPFAARLNTPMGPGRVTVVKEVNAVNAKSFNVDDLAKSKDIALHLNPLNIKAF 180  
QY 57 DRTLAWISRWGQKKLISAPFLFYPPQFFVLLFOEGGLKALNGQGLGATSNMQOALEQ 116  
Db 181 VRNSFLQESWGEEERNITSPFSPGMYFEMIIYCDVREFKAVNGVHSLEYKHFRELSS 240  
QY 117 LRELIRSGSVQLYCVHS 133  
Db 241 IDTLEINGDIHLLEVR 257  
RESULT 24  
AAY87403  
ID AAY87403 standard; protein; 316 AA.  
XX AC AAY87403;  
XX DT 03-JUL-2000 (first entry)  
XX DE Human PCTA-1 splice variant #1, SEQ ID NO:5.  
XX KW Prostate carcinoma antigen-1; PCTA-1; human; galectin-8 homologue;

KW splice variant; prostate cancer; biallelic marker; polymorphism;  
KW detection; diagnosis; prognosis; drug screening.  
OS Homo sapiens.  
XX WO9964590-A1.  
XX 16-DEC-1999.  
XX 04-JUN-1999; 99WO-IB001072.  
XX 05-JUN-1998; 98US-0088187P.  
PR 28-SEP-1998; 98US-0102324P.  
XX (GEST ) GENSET.  
XX Blumenfeld M, Bougueleret L, Chumakov I;  
XX WPI; 2000-338635/29.  
DR N-PSDB; AAA10226.  
XX Novel polynucleotide used to develop products for detecting  
PT susceptibility to, diagnosis, prognosis and therapy of prostate cancer.  
PT Claim 31; Page 329-330; 339pp; English.  
XX This sequence represents a splice variant (#1) of human prostate  
XX carcinoma tumour antigen-1 (PCTA-1). At least three PCTA-1 cDNAs  
CC (AAA10226, AAA10227, AAA10228) have been identified; the cDNA encoding  
CC this sequence comprises exons 0, 1, 2, 3, 4, 5, 6, 7, 8 and 9 out of the  
CC thirteen exons of the genomic PCTA-1 sequence (AAA10225). PCTA-1 is a  
CC human homologue of rat galectin-8, which can mediate both cell-cell and  
CC cell-matrix interactions. 125 biallelic markers were identified in the  
CC genomic sequence, with 40 biallelic markers being located in exonic  
CC regions. Six of these alter the amino acid sequence of a PCTA-1 protein.  
CC 43 markers were found in the 5' regulatory region, including 16 in the  
CC promoter, 39 in intronic sequences, and 3 in the 3' regulatory region.  
CC Alleles containing certain biallelic markers are associated with prostate  
CC cancer. These polymorphisms can be used as the basis of methods for  
CC determining genetic predisposition to prostate cancer, for diagnosis of  
CC this cancer and for the prognosis/detection of an eventual treatment  
CC response to therapeutic agents against prostate cancer. They can be used  
CC to design of drug screening protocols to provide accurate and efficient  
CC evaluation of therapeutic and side-effect potential of new or already  
CC existing medicaments, as well as for screening molecules which modulate  
CC or inhibit the expression of the PCTA-1 gene, preferably those active  
CC against prostate cancer  
XX

Query Match 16.1%; Score 109.5; DB 3; Length 316;  
Best Local Similarity 25.5%; Pred. No. 0.00011;  
Matches 35; Conservative 23; Mismatches 74; Indels 5; Gaps 1;  
QY 2 SPRLEVPCHALPQGLSPGVIIIRGLVLPQPKHPTVSL-----RDQAHAAPVTLRASFA 56  
DB 179 TPQLSLPFAARLNTPMGRTVVVKGEVNAKSFNVDDLAKGSKDIALHLNPLNIKAF 238  
QY 57 DRTLAWISRWQKKLISAPFLFYQRFVEVLLLFQEGGLKALNGQGLGATSMNQAALEQ 116  
DB 239 VRNSFLOESWGEEERNITSPFPSPGMYFEMIIYCDVREFKVAVNGVHSLKYKHKRFLSS 298  
QY 117 LRELIRISGSVOLYCVHS 133  
DB 299 IDTLEINGDIHLLVEVS 315  
RESULT 25  
AAB85030  
ID AAB85030 standard; protein; 316 AA.  
XX  
AC AAB85030;  
XX

06-AUG-2001 (first entry)  
Protein encoded by a PCTA-1 cDNA consisting of exons 0 to 9.  
BAP28; prostate; tumour; cancer; diagnostic; genetic analysis; PCTA-1.  
Homo sapiens.  
XX WO200100669-A2.  
XX 04-JAN-2001.  
XX 23-JUN-2000; 2000WO-IB001183.  
XX 25-JUN-1999; 99US-0141323P.  
PR 18-JAN-2000; 2000US-0176880P.  
XX (GEST ) GENSET.  
XX Barry C, Bougueleret L, Chumakov I, Cohen-Akenine A;  
XX WPI; 2001-367032/38.  
DR N-PSDB; AAP83912.  
XX New BAP28 polynucleotides and polypeptides overexpressed in prostate  
PT cancer cells for diagnosing prostate tumors, e.g. by hybridization or  
PT polymerase chain reaction assays.  
XX Example; Page 304-307; 349pp; English.  
XX The invention is directed to BAP28 polypeptides, BAP28 polynucleotide  
CC sequences and regulatory region located at the 3' and 5' ends of the  
CC BAP28 coding region. The BAP28 polypeptides can be expressed by standard  
CC recombinant methodology. BAP28 polynucleotides and polypeptides have been  
CC found to be over expressed in prostate tumour cells, therefore levels of  
CC BAP28 expression and/or activity may be assayed (e.g. by polymerase chain  
CC reaction (PCR)) to diagnose patient suffering from or susceptible to  
CC prostate cancer. Antibodies specific for the BAP28 polypeptides are  
CC useful as diagnostic reagents. Biallelic markers of the BAP28 gene are  
CC useful in genetic analysis. The present sequence represents a protein  
CC encoded by a first cDNA sequence of the PCTA-1 gene consisting of the  
CC exons 0 to 9. The coding strand of PCTA-1 gene is on the opposite of the  
CC coding strand of BAP28  
XX Sequence 316 AA;  
Query Match 16.1%; Score 109.5; DB 4; Length 316;  
Best Local Similarity 25.5%; Pred. No. 0.00011;  
Matches 35; Conservative 23; Mismatches 74; Indels 5; Gaps 1;  
QY 2 SPRLEVPCHALPQGLSPGVIIIRGLVLPQPKHPTVSL-----RDQAHAAPVTLRASFA 56  
DB 179 TPQLSLPFAARLNTPMGRTVVVKGEVNAKSFNVDDLAKGSKDIALHLNPLNIKAF 238  
QY 57 DRTLAWISRWQKKLISAPFLFYQRFVEVLLLFQEGGLKALNGQGLGATSMNQAALEQ 116  
DB 239 VRNSFLOESWGEEERNITSPFPSPGMYFEMIIYCDVREFKVAVNGVHSLKYKHKRFLSS 298  
QY 117 LRELIRISGSVOLYCVHS 133  
DB 299 IDTLEINGDIHLLVEVS 315  
RESULT 26  
ADOS7924  
ID ADOS7924 standard; protein; 316 AA.  
XX  
AC ADOS7924;  
XX  
DT 12-AUG-2004 (first entry)  
XX Human galectin-8 polypeptide.  
DE  
XX



CC or mRNA  
XX Sequence 317 AA;  
SQ  
Query Match 16.1%; Score 109.5; DB 2; Length 317;  
Best Local Similarity 25.5%; Pred. No. 0.00011;  
Matches 35; Conservative 23; Mismatches 74; Indels 5; Gaps 1;  
QY 2 SPRLVPCSHALPQGLSPGQVIVRGVLQEPKHPTVSL-----RDOAAHAPVTLRASFA 56  
Db 180 TPQLSLPFAARLNTPMGPGRTVVVKGVEVNAKSFVNDLLAGSKDIALHLNPLNIKAF 239  
QY 57 DRTLAWISRWGQKLLISAPFLFPQREFVLLLFQEGGLKALNGOGLGATSMNQALQ 116  
Db 240 VRNSFLOESWGEERNITAPFPSPGMVFEMIIYCDVREFKVAVNGVHSLYKHKRFKELSS 299  
QY 117 LRELISGSVOLYCVHS 133  
Db 300 IDTLEINGDIHLLEVR 316  
RESULT 28  
ADS34896  
ID ADS34896 standard; protein; 317 AA.  
AC ADS34896;  
DT 16-DEC-2004 (first entry)  
XX Human autoimmune disease-related protein - SEQ ID 110.  
DE  
KW single nucleotide polymorphism detection; SNP detection;  
KW rheumatoid arthritis; type 1 diabetes; multiple sclerosis;  
KW systemic lupus erythematosus; inflammatory bowel disease; psoriasis;  
KW thyroiditis; celiac disease; pernicious anaemia; asthma; vitiligo;  
KW glomerulonephritis; Grave's disease; myocarditis; Sjogren's disease;  
KW primary systemic vasculitis.  
XX Homo sapiens.  
XX WO2004083403-A2.  
XX 30-SEP-2004.  
XX 18-MAR-2004; 2004WO-US008461.  
XX 18-MAR-2003; 2003US-0455444P.  
XX 25-APR-2003; 2003US-0465241P.  
XX (APPL-) APPLERA CORP.  
XX Cargill M, Begovich AB, Alexander HC;  
XX WPI; 2004-728480/71.  
XX N-PSDB; ADS34813.  
XX New isolated nucleic acid molecule comprises at least 8 contiguous  
XX nucleotides where one of the nucleotides is a single nucleotide  
XX polymorphism (SNP), useful for diagnosing or treating autoimmune  
XX diseases, e.g. rheumatoid arthritis.  
XX Claim 12; SEQ ID NO 110; 123pp; English.  
XX The invention comprises amino acid and coding sequences containing  
XX genetic polymorphisms associated with an altered risk of developing an  
XX autoimmune disease (e.g. rheumatoid arthritis). The invention further  
XX comprises a method of identifying an individual that has an altered risk  
XX of developing an autoimmune disease, comprising detecting a single  
XX nucleotide polymorphism (SNP) in a nucleic acid of the invention. The DNA  
XX and protein sequences of the invention are useful for diagnosing and  
XX treating autoimmune diseases, such as: rheumatoid arthritis, type 1  
XX diabetes, multiple sclerosis, systemic lupus erythematosus, inflammatory  
XX bowel diseases, psoriasis, thyroiditis, celiac disease, pernicious

CC anaemia, asthma, vitiligo, glomerulonephritis, Grave's disease,  
CC myocarditis, Sjogren's disease, or primary systemic vasculitis. The  
CC present amino acid sequence represents a human autoimmune disease-related  
CC protein of the invention. NOTE: The present sequence is not shown in the  
CC specification, but has been retrieved from the WIPO website.  
XX Sequence 317 AA;  
SQ  
Query Match 16.1%; Score 109.5; DB 8; Length 317;  
Best Local Similarity 25.5%; Pred. No. 0.00011;  
Matches 35; Conservative 23; Mismatches 74; Indels 5; Gaps 1;  
QY 2 SPRLVPCSHALPQGLSPGQVIVRGVLQEPKHPTVSL-----RDOAAHAPVTLRASFA 56  
Db 180 TPQLRUPFAARLNTPMGPGRTVVVKGVEVNAKSFVNDLLAGSKDIALHLNPLNIKAF 239  
QY 57 DRTLAWISRWGQKLLISAPFLFPQREFVLLLFQEGGLKALNGOGLGATSMNQALQ 116  
Db 240 VRNSFLOESWGEERNITSPFPSPGMVFEMIIYCDVREFKVAVNGVHSLYKHKRFKELSS 299  
QY 117 LRELISGSVOLYCVHS 133  
Db 300 IDTLEINGDIHLLEVR 316  
RESULT 29  
ADS34895  
ID ADS34895 standard; protein; 317 AA.  
AC ADS34895;  
DT 16-DEC-2004 (first entry)  
XX Human autoimmune disease-related protein - SEQ ID 109.  
DE  
KW single nucleotide polymorphism detection; SNP detection;  
KW rheumatoid arthritis; type 1 diabetes; multiple sclerosis;  
KW systemic lupus erythematosus; inflammatory bowel disease; psoriasis;  
KW thyroiditis; celiac disease; pernicious anaemia; asthma; vitiligo;  
KW glomerulonephritis; Grave's disease; myocarditis; Sjogren's disease;  
KW primary systemic vasculitis.  
XX Homo sapiens.  
XX WO2004083403-A2.  
XX 30-SEP-2004.  
XX 18-MAR-2004; 2004WO-US008461.  
XX 18-MAR-2003; 2003US-0455444P.  
XX 25-APR-2003; 2003US-0465241P.  
XX (APPL-) APPLERA CORP.  
XX Cargill M, Begovich AB, Alexander HC;  
XX WPI; 2004-728480/71.  
XX N-PSDB; ADS34812.  
XX New isolated nucleic acid molecule comprises at least 8 contiguous  
XX nucleotides where one of the nucleotides is a single nucleotide  
XX polymorphism (SNP), useful for diagnosing or treating autoimmune  
XX diseases, e.g. rheumatoid arthritis.  
XX Claim 12; SEQ ID NO 109; 123pp; English.  
XX The invention comprises amino acid and coding sequences containing  
XX genetic polymorphisms associated with an altered risk of developing an  
XX autoimmune disease (e.g. rheumatoid arthritis). The invention further  
XX comprises a method of identifying an individual that has an altered risk  
XX of developing an autoimmune disease, comprising detecting a single  
XX nucleotide polymorphism (SNP) in a nucleic acid of the invention. The DNA  
XX and protein sequences of the invention are useful for diagnosing and  
XX treating autoimmune diseases, such as: rheumatoid arthritis, type 1  
XX diabetes, multiple sclerosis, systemic lupus erythematosus, inflammatory  
XX bowel diseases, psoriasis, thyroiditis, celiac disease, pernicious



CC and protein sequences of the invention are useful for diagnosing and  
CC treating autoimmune diseases, such as: rheumatoid arthritis, type 1  
CC diabetes, multiple sclerosis, systemic lupus erythematosus, inflammatory  
CC bowel diseases, psoriasis, thyroiditis, celiac disease, pernicious  
CC anaemia, asthma, vitiligo, glomerulonephritis, Grave's disease,  
CC myocarditis, Sjogren's disease, or primary systemic vasculitis. The  
CC present amino acid sequence represents a human autoimmune disease-related  
CC protein of the invention. NOTE: The present sequence is not shown in the  
CC specification, but has been retrieved from the WIPO website.  
XX  
SQ Sequence 317 AA;  
  
Query Match 16.1%; Score 109.5; DB 8; Length 317;  
Best Local Similarity 25.5%; Pred. No. 0.00011;  
Matches 35; Conservative 23; Mismatches 74; Indels 5; Gaps 1;  
  
QY 2 SPRLEVPCHALPQGLSPQGVIIIRGLVLQEPKHPTVSL-----RDQAAHAPVTLRASFA 56  
Db 180 TPQLRLPFAARLNTPMGPGRTVVVKGEVANAKSFNVDLGKSKDIALHLNPLNIKAF 239  
  
QY 57 DRTLAWISRWGQKKLISAPFLYPQRPFFVLLFQGGKLLALNGQGLGATSMNQQALEQ 116  
Db 240 VRNSFLQESWGEERNITSPFSPGMYFEMIIYCDVREFKAVNGVHSLKYHRFKELSS 299  
  
QY 117 LRELISGVSQLYCVHS 133  
Db 300 IDTLEINGDIHLEVR 316  
  
RESULT 30  
ABP65200  
ID ABP65200 standard; protein; 318 AA.  
XX  
AC ABP65200;  
XX  
DT 12-NOV-2002 (first entry)  
XX  
DE Hypoxia-regulated protein #74.  
XX  
KW Cytostatic; vasotropic; tranquiliser; antiatherosclerotic; gene therapy;  
KW antiinflammatory; vulnery; gynecological; ophthalmological; vaccine;  
KW hypoxia; tumorigenesis; angiogenesis; apoptosis; cancer;  
KW ischaemic condition; reperfusion injury; retinopathy; neonatal stress;  
KW preeclampsia; atherosclerosis; inflammatory condition; wound healing;  
KW inflammation; erythropoiesis; hair loss; human.  
XX  
OS Homo sapiens.  
XX  
PN WO200246465-A2.  
XX  
PD 13-JUN-2002.  
XX  
PF 10-DEC-2001; 2001WO-GB005458.  
XX  
PR 08-DEC-2000; 2000GB-00030076.  
PR 08-FEB-2001; 2001GB-00003156.  
PR 25-OCT-2001; 2001GB-00025666.  
XX  
PA (OXFO-) OXFORD BIOMEDICA UK LTD.  
XX  
PI White J, Mundy CR, Ward NR, Krige D, Kingsman SM, Harris RA;  
PI Rayner WN;  
XX  
DR WPI; 2002-627238/67.  
XX  
PT Identifying a gene involved in disease for treating hypoxia-regulated  
PT conditions, comprises comparing the transcriptome/proteome of two cell  
PT types under different conditions and identifying a differentially  
PT regulated gene.  
XX  
PS Claim 35; Page 389; 538pp; English.  
XX  
CC The present invention relates to methods for identifying genes and

CC proteins that are implicated in a specific disease or physiological  
CC condition. The method comprises comparing the transcriptome/proteome of a  
CC specialised cell type implicated in a disease or condition with that of a  
CC second specialised cell type, under two experimental conditions, and  
CC identifying a gene that is differentially regulated in the two  
CC specialised cell types under experimental conditions. ABV7873-ABV78116  
CC and ABP65061-ABP65257 were identified using the methods of the invention.  
CC The coding sequences and proteins are useful for treating a disease in a  
CC patient, for manufacture of a medicament for treating hypoxia-regulated  
CC conditions, and for regulating tumorigenesis, angiogenesis, apoptosis,  
CC biological response to hypoxia conditions, or hypoxic-associated  
CC pathology in a patient. The coding sequences and proteins are also useful  
CC for monitoring the therapeutic treatment of a disease or physiological  
CC condition, such as cancer, ischaemic conditions, reperfusion injury,  
CC retinopathy, neonatal stress, preeclampsia, atherosclerosis, inflammatory  
CC conditions, wound healing, inflammation, erythropoiesis or hair loss  
XX  
SQ Sequence 318 AA;  
  
Query Match 16.1%; Score 109.5; DB 5; Length 318;  
Best Local Similarity 25.5%; Pred. No. 0.00011;  
Matches 35; Conservative 23; Mismatches 74; Indels 5; Gaps 1;  
  
QY 2 SPRLEVPCHALPQGLSPQGVIIIRGLVLQEPKHPTVSL-----RDQAAHAPVTLRASFA 56  
Db 181 TPQLRLPFAARLNTPMGPGRTVVVKGEVANAKSFNVDLGKSKDIALHLNPLNIKAF 240  
  
QY 57 DRTLAWISRWGQKKLISAPFLYPQRPFFVLLFQGGKLLALNGQGLGATSMNQQALEQ 116  
Db 241 VRNSFLQESWGEERNITSPFSPGMYFEMIIYCDVREFKAVNGVHSLKYHRFKELSS 300  
  
QY 117 LRELISGVSQLYCVHS 133  
Db 301 IDTLEINGDIHLEVR 317  
  
Search completed: October 17, 2005, 08:39:25  
Job time : 170 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2005, 08:35:58 ; Search time 42 Seconds  
(without alignments)  
236.389 Million cell updates/sec

Title: HJA54  
Perfect score: 682  
Sequence: 1 msrlevpcshapqglsps.....leqlrelrsgsvglycvhs 133

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

- Issued Patents AA.\*  
1: /cgm2\_6/ptodata/1/iaa/5A COMB.pep.\*  
2: /cgm2\_6/ptodata/1/iaa/5B COMB.pep.\*  
3: /cgm2\_6/ptodata/1/iaa/6A COMB.pep.\*  
4: /cgm2\_6/ptodata/1/iaa/6B COMB.pep.\*  
5: /cgm2\_6/ptodata/1/iaa/PTCUS COMB.pep.\*  
6: /cgm2\_6/ptodata/1/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	671	98.4	133	4	US-09-557-170A-2
2	671	98.4	275	4	US-09-557-170A-25
3	671	98.4	296	4	US-09-557-170A-27
4	671	98.4	336	3	US-09-131-648-1
5	119	17.4	145	2	US-08-788-584-5
6	119	17.4	145	3	US-08-946-914-12
7	119	17.4	145	4	US-09-656-450-12
8	119	17.4	145	4	US-09-557-170A-3
9	109.5	16.1	316	4	US-09-326-402C-5
10	109.5	16.1	317	3	US-09-326-402C-15
11	109.5	16.1	317	4	US-08-946-914-6
12	109.5	16.1	317	4	US-09-656-450-6
13	109.5	16.1	318	4	US-09-557-170A-4
14	108.5	15.9	316	3	US-09-131-648-5
15	108.5	15.9	317	4	US-08-875-553D-30
16	106.5	15.6	368	4	US-09-326-402C-7
17	98	14.4	322	4	US-09-559-023-2
18	96.5	14.1	358	4	US-09-326-402C-6
19	96.5	14.1	358	4	US-09-326-402C-16
20	95.5	14.0	135	4	US-09-326-402C-20
21	95.5	14.0	136	3	US-08-946-914-13
22	95.5	14.0	136	4	US-09-154-750A-79
23	95.5	14.0	136	4	US-09-656-450-13
24	94	13.8	315	4	US-09-326-402C-17
25	91.5	13.4	316	4	US-09-326-402C-9
26	90.5	13.3	139	4	US-09-513-999C-6729
27	88.5	13.0	264	1	US-08-562-311-4
Sequence 2, Appli					
Sequence 25, Appl					
Sequence 27, Appl					
Sequence 1, Appli					
Sequence 5, Appli					
Sequence 12, Appl					
Sequence 12, Appl					
Sequence 12, Appl					
Sequence 3, Appli					
Sequence 5, Appli					
Sequence 15, Appl					
Sequence 6, Appli					
Sequence 6, Appli					
Sequence 4, Appli					
Sequence 5, Appli					
Sequence 30, Appl					
Sequence 7, Appli					
Sequence 2, Appli					
Sequence 6, Appli					
Sequence 16, Appl					
Sequence 20, Appl					
Sequence 13, Appl					
Sequence 17, Appl					
Sequence 9, Appli					
Sequence 4, Appli					

28	87.5	12.8	316	2	US-08-728-521-3	Sequence 3, Appli
29	87.5	12.8	316	2	US-08-647-960-2	Sequence 2, Appli
30	87.5	12.8	316	2	US-08-946-914-15	Sequence 15, Appl
31	87.5	12.8	316	3	US-08-946-914-17	Sequence 3, Appli
32	87.5	12.8	316	3	US-09-212-146-3	Sequence 15, Appl
33	87.5	12.8	316	4	US-09-656-450-15	Sequence 17, Appl
34	87.5	12.8	316	4	US-09-656-450-17	Sequence 1, Appli
35	86.5	12.7	264	2	US-08-728-521-1	Sequence 11, Appli
36	86.5	12.7	264	3	US-09-212-146-1	Sequence 11, Appli
37	84	12.3	324	3	US-08-946-914-11	Sequence 2, Appli
38	84	12.3	324	4	US-09-656-450-11	Sequence 11, Appl
39	82.5	12.1	250	1	US-08-562-311-2	Sequence 2, Appli
40	81.5	12.0	323	1	US-08-469-667-16	Sequence 16, Appl
41	81.5	12.0	323	3	US-08-946-914-2	Sequence 2, Appli
42	81.5	12.0	323	3	US-09-224-110-16	Sequence 16, Appl
43	81.5	12.0	323	4	US-09-656-450-2	Sequence 2, Appli
44	81.5	12.0	323	4	US-09-326-402C-21	Sequence 21, Appl
45	81.5	12.0	323	4	US-09-988-292A-16	Sequence 16, Appl
46	81.5	12.0	323	5	PCT-US95-07289-16	Sequence 15, Appl
47	81.5	12.0	341	4	US-09-949-016-7251	Sequence 7251, Ap
48	81	11.9	273	4	US-09-902-540-9846	Sequence 9846, Ap
49	81	11.9	737	3	US-09-511-625B-2	Sequence 2, Appli
50	81	11.9	819	3	US-09-511-625B-4	Sequence 4, Appli
51	81	11.9	847	1	US-08-276-099A-2	Sequence 2, Appli
52	81	11.9	847	1	US-08-781-890-2	Sequence 2, Appli
53	81	11.9	847	3	US-09-087-465-12	Sequence 12, Appl
54	81	11.9	847	3	US-09-511-625B-68	Sequence 68, Appl
55	81	11.9	847	4	US-09-972-800A-10	Sequence 10, Appl
56	80	11.7	145	2	US-08-788-584-1	Sequence 1, Appli
57	80	11.7	200	3	US-08-946-914-8	Sequence 8, Appli
58	80	11.7	200	4	US-09-656-450-8	Sequence 8, Appli
59	78.5	11.5	262	3	US-08-946-914-14	Sequence 14, Appl
60	78.5	11.5	262	4	US-09-656-450-14	Sequence 14, Appl
61	78.5	11.5	282	4	US-09-489-039A-12241	Sequence 12241, A
62	77.5	11.4	143	4	US-09-877-790-1	Sequence 1, Appli
63	77.5	11.4	249	4	US-09-538-092-951	Sequence 951, App
64	77.5	11.4	250	3	US-08-946-914-10	Sequence 10, Appl
65	77.5	11.4	250	4	US-09-656-450-10	Sequence 10, Appl
66	77.5	11.4	250	4	US-09-919-039-298	Sequence 298, App
67	77.5	11.4	250	4	US-09-877-790-2	Sequence 2, Appli
68	77.5	11.4	251	4	US-09-949-016-7560	Sequence 7560, Ap
69	76	11.1	311	3	US-08-946-914-4	Sequence 4, Appli
70	76	11.1	311	4	US-09-656-450-4	Sequence 4, Appli
71	76	11.1	355	4	US-09-326-402C-18	Sequence 18, Appl
72	76	11.1	378	4	US-09-854-133-439	Sequence 439, App
73	75	11.0	149	2	US-08-788-584-3	Sequence 3, Appli
74	75	11.0	343	4	US-09-252-991A-27631	Sequence 27631, A
75	75	11.0	368	4	US-09-902-540-13302	Sequence 13302, A
76	75	11.0	369	4	US-09-543-681A-7514	Sequence 7514, Ap
77	74.5	10.9	146	2	US-08-647-960-8	Sequence 8, Appli
78	73	10.7	323	4	US-09-326-402C-19	Sequence 19, Appl
79	71.5	10.5	332	4	US-09-135-121B-7	Sequence 7, Appli
80	71	10.4	134	4	US-09-326-402C-14	Sequence 14, Appl
81	71	10.4	135	3	US-08-946-914-16	Sequence 16, Appl
82	71	10.4	135	3	US-08-050-459B-20	Sequence 20, Appl
83	71	10.4	135	3	US-09-489-232-2	Sequence 2, Appli
84	71	10.4	135	4	US-08-719-579-2	Sequence 2, Appli
85	71	10.4	135	4	US-09-656-450-16	Sequence 16, Appl
86	71	10.4	135	4	US-09-919-172-91	Sequence 91, Appl
87	71	10.4	135	4	US-09-976-594-628	Sequence 628, App
88	71	10.4	135	4	US-09-919-039-242	Sequence 242, App
89	71	10.4	135	4	US-09-919-497-81	Sequence 81, Appl
90	71	10.4	149	4	US-09-949-016-10945	Sequence 10945, A
91	71	10.4	422	4	US-09-270-767-42572	Sequence 42572, A
92	71	10.4	617	4	US-09-252-991A-21113	Sequence 21113, A
93	70	10.3	303	4	US-09-583-110-4995	Sequence 4995, Ap
94	70	10.3	347	4	US-09-107-433-5171	Sequence 5171, Ap
95	70	10.3	373	4	US-09-134-000C-5355	Sequence 5355, Ap
96	70	10.3	545	4	US-09-949-016-10223	Sequence 10223, A
97	69.5	10.2	206	4	US-09-180-109A-15	Sequence 15, Appl
98	69.5	10.2	206	4	US-09-180-109A-18	Sequence 18, Appl
99	69	10.1	352	3	US-09-134-001C-4121	Sequence 4121, Ap
100	68.5	10.0	320	4	US-09-252-991A-23385	Sequence 23385, A

## ALIGNMENTS

```
RESULT 1
US-09-557-170A-2
; Sequence 2, Application US/09557170A
; Patent No. 6605699
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: Galectin 11
; FILE REFERENCE: PF354P2
; CURRENT APPLICATION NUMBER: US/09/557,170A
; CURRENT FILING DATE: 2000-04-21
; PRIOR APPLICATION NUMBER: 09/109,864
; PRIOR FILING DATE: 1998-06-06
; PRIOR APPLICATION NUMBER: 09/010,146
; PRIOR FILING DATE: 1998-01-21
; PRIOR APPLICATION NUMBER: 60/034,205
; PRIOR FILING DATE: 1997-01-21
; PRIOR APPLICATION NUMBER: 60/034,204
; PRIOR FILING DATE: 1997-01-21
; PRIOR APPLICATION NUMBER: 60/169,932
; PRIOR FILING DATE: 1999-12-10
; PRIOR APPLICATION NUMBER: 60/130,390
; PRIOR FILING DATE: 1999-04-21
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 133
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-557-170A-2

Query Match          98.4%; Score 671; DB 4; Length 133;
Best Local Similarity 99.2%; Pred. No. 1.3e-76;
Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSPRLVPCSHALPQGLSPGQVIIVRGLVLOEPKHPFTVSLRDOAAHAPVTLRASPADRTL 60
Db 1 MSPRLVPCSHALPQGLSPGQVIIVRGLVLOEPKHPFTVSLRDOAAHAPVTLRASPADRTL 60
QY 61 AWISRWGQKKLIASAPFLFYPPQRFVFFVLLLFQEGGLKALNGQGLGATSMNQQALEQLREL 120
Db 61 AWISRWGQKKLIASAPFLFYPPQRFVFFVLLLFQEGGLKALNGQGLGATSMNQQALEQLREL 120
QY 121 RISGSVOLYCVHS 133
Db 121 RISGSVOLYCVHS 133

RESULT 2
US-09-557-170A-25
; Sequence 25, Application US/09557170A
; Patent No. 6605699
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: Galectin 11
; FILE REFERENCE: PF354P2
; CURRENT APPLICATION NUMBER: US/09/557,170A
; CURRENT FILING DATE: 2000-04-21
; PRIOR APPLICATION NUMBER: 09/109,864
; PRIOR FILING DATE: 1998-06-06
; PRIOR APPLICATION NUMBER: 09/010,146
; PRIOR FILING DATE: 1998-01-21
; PRIOR APPLICATION NUMBER: 60/034,205
; PRIOR FILING DATE: 1997-01-21
; PRIOR APPLICATION NUMBER: 60/034,204
; PRIOR FILING DATE: 1997-01-21
; PRIOR APPLICATION NUMBER: 60/169,932
; PRIOR FILING DATE: 1999-12-10
; PRIOR APPLICATION NUMBER: 60/130,390
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-557-170A-27

Query Match          98.4%; Score 671; DB 4; Length 296;
Best Local Similarity 99.2%; Pred. No. 3.9e-76;
Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSPRLVPCSHALPQGLSPGQVIIVRGLVLOEPKHPFTVSLRDOAAHAPVTLRASPADRTL 60
Db 164 MSPRLVPCSHALPQGLSPGQVIIVRGLVLOEPKHPFTVSLRDOAAHAPVTLRASPADRTL 223
QY 61 AWISRWGQKKLIASAPFLFYPPQRFVFFVLLLFQEGGLKALNGQGLGATSMNQQALEQLREL 120
Db 224 AWISRWGQKKLIASAPFLFYPPQRFVFFVLLLFQEGGLKALNGQGLGATSMNQQALEQLREL 283
QY 121 RISGSVOLYCVHS 133
Db 284 RISGSVOLYCVHS 296
```

```
; PRIOR FILING DATE: 1999-04-21
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-557-170A-25
```

```
Query Match          98.4%; Score 671; DB 4; Length 275;
Best Local Similarity 99.2%; Pred. No. 3.5e-76;
Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSPRLVPCSHALPQGLSPGQVIIVRGLVLOEPKHPFTVSLRDOAAHAPVTLRASPADRTL 60
Db 143 MSPRLVPCSHALPQGLSPGQVIIVRGLVLOEPKHPFTVSLRDOAAHAPVTLRASPADRTL 202
QY 61 AWISRWGQKKLIASAPFLFYPPQRFVFFVLLLFQEGGLKALNGQGLGATSMNQQALEQLREL 120
Db 203 AWISRWGQKKLIASAPFLFYPPQRFVFFVLLLFQEGGLKALNGQGLGATSMNQQALEQLREL 262
QY 121 RISGSVOLYCVHS 133
Db 263 RISGSVOLYCVHS 275
```

```
RESULT 3
US-09-557-170A-27
; Sequence 27, Application US/09557170A
; Patent No. 6605699
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: Galectin 11
; FILE REFERENCE: PF354P2
; CURRENT APPLICATION NUMBER: US/09/557,170A
; CURRENT FILING DATE: 2000-04-21
; PRIOR APPLICATION NUMBER: 09/109,864
; PRIOR FILING DATE: 1998-06-06
; PRIOR APPLICATION NUMBER: 09/010,146
; PRIOR FILING DATE: 1998-01-21
; PRIOR APPLICATION NUMBER: 60/034,205
; PRIOR FILING DATE: 1997-01-21
; PRIOR APPLICATION NUMBER: 60/034,204
; PRIOR FILING DATE: 1997-01-21
; PRIOR APPLICATION NUMBER: 60/169,932
; PRIOR FILING DATE: 1999-12-10
; PRIOR APPLICATION NUMBER: 60/130,390
; PRIOR FILING DATE: 1999-04-21
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-557-170A-27
```

```
Query Match          98.4%; Score 671; DB 4; Length 296;
Best Local Similarity 99.2%; Pred. No. 3.9e-76;
Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSPRLVPCSHALPQGLSPGQVIIVRGLVLOEPKHPFTVSLRDOAAHAPVTLRASPADRTL 60
Db 164 MSPRLVPCSHALPQGLSPGQVIIVRGLVLOEPKHPFTVSLRDOAAHAPVTLRASPADRTL 223
QY 61 AWISRWGQKKLIASAPFLFYPPQRFVFFVLLLFQEGGLKALNGQGLGATSMNQQALEQLREL 120
Db 224 AWISRWGQKKLIASAPFLFYPPQRFVFFVLLLFQEGGLKALNGQGLGATSMNQQALEQLREL 283
QY 121 RISGSVOLYCVHS 133
Db 284 RISGSVOLYCVHS 296
```



```
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-946-914-12

Query Match      17.4%; Score 119; DB 3; Length 145;
Best Local Similarity 26.9%; Pred. No. 5.6e-07;
Matches 35; Conservative 26; Mismatches 65; Indels 4; Gaps 2;

QY 3 PRLEVPCHALPGQLSPGQVVIIVRGLVLOEPKHPHTVSLR---DQAAHAPVTLRASPADRT 59
Db 11 PNLAVPFFTSIPNGLYPSKSIIVISGVVLSDAKRFQINLRCGGDIAFHLNPRFDENAVVRN 70
QY 60 LAWISRWG-QKKLISAPFLFYQRPFEVLLLFQEGGLKALNGQGLGATSMNQALEQLR 118
Db 71 TQINNSWGPEERSLPGSMPPFSRGQRFVSWILCEGHCFKVAVDGQHICEYSHRLMNLDPIN 130
QY 119 ELRISGSVOL 128
Db 131 TLEVAGDIQL 140

RESULT 7
US-09-656-450-12
; Sequence 12, Application US/09656450
; Patent No. 6468768
; GENERAL INFORMATION:
; APPLICANT: Ni, Jian
; APPLICANT: Gentz, Reiner L.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Galectin 9 and 10SV Polynucleotides
; FILE REFERENCE: 1488.0560003
; CURRENT APPLICATION NUMBER: US/09/656,450
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: US 09/263,689
; PRIOR FILING DATE: 1999-03-05
; PRIOR APPLICATION NUMBER: US 08/946,914
; PRIOR FILING DATE: 1997-10-09
; PRIOR APPLICATION NUMBER: US 60/028,093
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 12
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Rat
US-09-656-450-12

Query Match      17.4%; Score 119; DB 4; Length 145;
Best Local Similarity 26.9%; Pred. No. 5.6e-07;
Matches 35; Conservative 26; Mismatches 65; Indels 4; Gaps 2;

QY 3 PRLEVPCHALPGQLSPGQVVIIVRGLVLOEPKHPHTVSLR---DQAAHAPVTLRASPADRT 59
Db 11 PNLAVPFFTSIPNGLYPSKSIIVISGVVLSDAKRFQINLRCGGDIAFHLNPRFDENAVVRN 70
QY 60 LAWISRWG-QKKLISAPFLFYQRPFEVLLLFQEGGLKALNGQGLGATSMNQALEQLR 118
Db 71 TQINNSWGPEERSLPGSMPPFSRGQRFVSWILCEGHCFKVAVDGQHICEYSHRLMNLDPIN 130
QY 119 ELRISGSVOL 128
Db 131 TLEVAGDIQL 140

RESULT 8
US-09-557-170A-3
; Sequence 3, Application US/09557170A
; Patent No. 6605699
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: Galectin 11
```

```
; FILE REFERENCE: PF354P2
; CURRENT APPLICATION NUMBER: US/09/557,170A
; CURRENT FILING DATE: 2000-04-21
; PRIOR APPLICATION NUMBER: 09/109,864
; PRIOR FILING DATE: 1998-06-06
; PRIOR APPLICATION NUMBER: 09/010,146
; PRIOR FILING DATE: 1998-01-21
; PRIOR APPLICATION NUMBER: 60/034,205
; PRIOR FILING DATE: 1997-01-21
; PRIOR APPLICATION NUMBER: 60/034,204
; PRIOR FILING DATE: 1997-01-21
; PRIOR APPLICATION NUMBER: 60/169,932
; PRIOR FILING DATE: 1999-12-10
; PRIOR APPLICATION NUMBER: 60/130,390
; PRIOR FILING DATE: 1999-04-21
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-557-170A-3

Query Match      17.4%; Score 119; DB 4; Length 145;
Best Local Similarity 26.9%; Pred. No. 5.6e-07;
Matches 35; Conservative 26; Mismatches 65; Indels 4; Gaps 2;

QY 3 PRLEVPCHALPGQLSPGQVVIIVRGLVLOEPKHPHTVSLR---DQAAHAPVTLRASPADRT 59
Db 11 PNLAVPFFTSIPNGLYPSKSIIVISGVVLSDAKRFQINLRCGGDIAFHLNPRFDENAVVRN 70
QY 60 LAWISRWG-QKKLISAPFLFYQRPFEVLLLFQEGGLKALNGQGLGATSMNQALEQLR 118
Db 71 TQINNSWGPEERSLPGSMPPFSRGQRFVSWILCEGHCFKVAVDGQHICEYSHRLMNLDPIN 130
QY 119 ELRISGSVOL 128
Db 131 TLEVAGDIQL 140

RESULT 9
US-09-326-402C-5
; Sequence 5, Application US/09326402C
; Patent No. 6759192
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; TITLE OF INVENTION: Polymorphic Markers of Prostate Carcinoma Tumor Antigen-1 (PCTA-1)
; FILE REFERENCE: GEN-T112XCI
; CURRENT APPLICATION NUMBER: US/09/326,402C
; CURRENT FILING DATE: 1999-06-04
; PRIOR APPLICATION NUMBER: 60/088,187
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/102,324
; PRIOR FILING DATE: 1998-09-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 5
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 18
; OTHER INFORMATION: 5-2-162 : polymorphic amino acid Tyr or Phe
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 35
; OTHER INFORMATION: 5-2-213 : polymorphic amino acid Cys or Arg
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 55
```

```
; OTHER INFORMATION: 5-3-84 : polymorphic amino acid Val or Met
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 183
; OTHER INFORMATION: 5-7-195 : polymorphic amino acid Ser or Arg
US-09-326-402C-5

Query Match      16.1%; Score 109.5; DB 4; Length 316;
Best Local Similarity 25.5%; Pred. No. 2.6e-05;
Matches 35; Conservative 23; Mismatches 74; Indels 5; Gaps 1;

QY 2 SPRLVPCSHALPQGLSPGQVIVIRGLVLPQPKHPTVSL-----RDQAAHAPVTLRASFA 56
Db 179 TPQLSLPFAARLNTPMGPGRTVVVKGEVNAKSFNVDLLAGKSKDIALHNLPRNLKAF 238
QY 57 DRTLAWISRWGQKKLISAPFLYPQRFPEVLLFQEGGLKALNGQGLGATSMNQALEQ 116
Db 239 VRNSFLOESWGEERNTSPFPSPGMYFEMIIYCDVREFKVAVNGVHSLEYKHKRFKELSS 298
QY 117 LRELIRISGSVOLYCVHS 133
Db 299 IDTLEINGDIHLLEVR 315

RESULT 10
US-09-326-402C-15
; Sequence 15, Application US/09326402C
; Patent No. 6759192
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; TITLE OF INVENTION: Polymorphic Markers of Prostate Carcinoma Tumor Antigen-1 (PCTA-1)
; FILE REFERENCE: GEN-TL2XC1
; CURRENT APPLICATION NUMBER: US/09/326,402C
; CURRENT FILING DATE: 1999-06-04
; PRIOR APPLICATION NUMBER: 60/088,187
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/102,324
; PRIOR FILING DATE: 1998-09-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 15
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(316)
; OTHER INFORMATION: amino acid sequence of PCTA
US-09-326-402C-15

Query Match      16.1%; Score 109.5; DB 4; Length 316;
Best Local Similarity 25.5%; Pred. No. 2.6e-05;
Matches 35; Conservative 23; Mismatches 74; Indels 5; Gaps 1;

QY 2 SPRLVPCSHALPQGLSPGQVIVIRGLVLPQPKHPTVSL-----RDQAAHAPVTLRASFA 56
Db 179 TPQLSLPFAARLNTPMGPGRTVVVKGEVNAKSFNVDLLAGKSKDIALHNLPRNLKAF 238
QY 57 DRTLAWISRWGQKKLISAPFLYPQRFPEVLLFQEGGLKALNGQGLGATSMNQALEQ 116
Db 239 VRNSFLOESWGEERNTSPFPSPGMYFEMIIYCDVREFKVAVNGVHSLEYKHKRFKELSS 298
QY 117 LRELIRISGSVOLYCVHS 133
Db 299 IDTLEINGDIHLLEVR 315

RESULT 11
US-08-946-914-6
; Sequence 6, Application US/08946914
; Patent No. 6027916
```

```
; GENERAL INFORMATION:
; APPLICANT: Ni, Jian
; APPLICANT: Gentz, Reiner L.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Galectin 8, 9, 10 and 10SV
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein, & Fox P.L.L.C.
; STREET: 1100 New York Ave., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/946,914
; FILING DATE: Herewith
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/028,093
; FILING DATE: 09-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Steffe, Eric K.
; REGISTRATION/DOCKET NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.0560001/EKS/SGW
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 317 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-946-914-6

Query Match      16.1%; Score 109.5; DB 3; Length 317;
Best Local Similarity 25.5%; Pred. No. 2.6e-05;
Matches 35; Conservative 23; Mismatches 74; Indels 5; Gaps 1;

QY 2 SPRLVPCSHALPQGLSPGQVIVIRGLVLPQPKHPTVSL-----RDQAAHAPVTLRASFA 56
Db 180 TPQLSLPFAARLNTPMGPGRTVVVKGEVNAKSFNVDLLAGKSKDIALHNLPRNLKAF 239
QY 57 DRTLAWISRWGQKKLISAPFLYPQRFPEVLLFQEGGLKALNGQGLGATSMNQALEQ 116
Db 240 VRNSFLOESWGEERNTAFPFSPGMYFEMIIYCDVREFKVAVNGVHSLEYKHKRFKELSS 299
QY 117 LRELIRISGSVOLYCVHS 133
Db 300 IDTLEINGDIHLLEVR 316

RESULT 12
US-09-656-450-6
; Sequence 6, Application US/09656450
; Patent No. 6468768
; GENERAL INFORMATION:
; APPLICANT: Ni, Jian
; APPLICANT: Gentz, Reiner L.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Galectin 9 and 10SV Polynucleotides
; FILE REFERENCE: 1488.0560003
; CURRENT APPLICATION NUMBER: US/09/656,450
; CURRENT FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: US 09/263,689
; PRIOR FILING DATE: 1999-03-05
; PRIOR APPLICATION NUMBER: US 08/946,914
; PRIOR FILING DATE: 1997-10-09
```











QY 4 RLEVPCHALPQGLSPQGVIIVRGLVLOEPKHTVSL-----RDQAAHAPVTLRASFADR 58  
Db 181 QLSLPEARLNASMGRTVWIKGEVNTNARSFVNDLVAGKTRDIALHLNPLNKAFV-R 239  
QY 59 TLAWISRWGQKKLISAPFLFYPPQRFVEVLLFOEGGLKALNQGGIGATSMNQALEQLR 118  
Db 240 NSFLQDANGEEERNITCFPPSSGMFYEMIIYCDVREFKVAINGVHSLEYKHKRFDLSSID 299  
QY 119 ELRISGSVQLYCVHS 133  
Db 300 TLSVDGDIRLLDVR 314  
RESULT 25  
US-09-326-402C-9  
; Sequence 9, Application US/09326402C  
; Patent No. 6759192  
; GENERAL INFORMATION:  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Bougueleret, Lydie  
; APPLICANT: Chumakov, Ilya  
; TITLE OF INVENTION: Polymorphic Markers of Prostate Carcinoma Tumor Antigen-1 (PCTA-1)  
; FILE REFERENCE: GEN-T112XC1  
; CURRENT APPLICATION NUMBER: US/09/326,402C  
; CURRENT FILING DATE: 1999-06-04  
; PRIOR APPLICATION NUMBER: 60/088,187  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/102,324  
; PRIOR FILING DATE: 1998-09-28  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 9  
; LENGTH: 316  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-326-402C-9  
Query Match 13.4%; Score 91.5; DB 4; Length 316;  
Best Local Similarity 23.0%; Pred. No. 0.0048;  
Matches 31; Conservative 23; Mismatches 76; Indels 5; Gaps 1;  
QY 4 RLEVPCHALPQGLSPQGVIIVRGLVLOEPKHTVSL-----RDQAAHAPVTLRASFADR 58  
Db 181 QLSLPEARLNASMGRTVWIKGEVNTNARSFVNDLVAGKTRDIALHLNPLNKAFV-R 240  
QY 59 TLAWISRWGQKKLISAPFLFYPPQRFVEVLLFOEGGLKALNQGGIGATSMNQALEQLR 118  
Db 241 NSFLQDANGEEERNITCFPPSSGMFYEMIIYCDVREFKVAINGVHSLEYKHKRFDLSSID 300  
QY 119 ELRISGSVQLYCVHS 133  
Db 301 TLSVDGDIRLLDVR 315  
RESULT 26  
US-09-513-999C-6729  
; Sequence 6729, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; FILE REFERENCE: 59.US2.REG  
; CURRENT APPLICATION NUMBER: US/09/513,999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pgm  
; SEQ ID NO 6729

; LENGTH: 139  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-513-999C-6729  
Query Match 13.3%; Score 90.5; DB 4; Length 139;  
Best Local Similarity 33.7%; Pred. No. 0.0021;  
Matches 31; Conservative 15; Mismatches 35; Indels 11; Gaps 5;  
QY 3 PRLVPCSHALPQGLSPQGVIIVRGLVLOEPKHTVSL-----RDQAAHAPVTLRASFA 56  
Db 34 PRLVPCFCHG-IKGGMRPGKVLVNGIVDLNPESFAISLTGDSDEPPADVAIELKAVFT 92  
QY 57 DRTL---AMIS-RWGOKLISAPFLFYPPQRF 84  
Db 93 DRQLLNSCISGERGEEQSAIFYFFIFPDQPF 124  
RESULT 27  
US-08-562-311-4  
; Sequence 4, Application US/08562311  
; Patent No. 5801002  
; GENERAL INFORMATION:  
; APPLICANT: RAZ, AVRAHAM  
; TITLE OF INVENTION: A METHOD OF DETERMINING THE PROBABILITY  
; TITLE OF INVENTION: OF METASTASIS IN A CELL SAMPLE  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dykema Gossett  
; STREET: STE 505 N. Woodward  
; CITY: Bloomfield Hills  
; STATE: MI  
; COUNTRY: U.S.  
; ZIP: 48304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/562,311  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/188,225  
; FILING DATE:  
; APPLICATION NUMBER: US 07/681,242  
; FILING DATE: 04-APR-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/294,249  
; FILING DATE: 01-JUN-1989  
; ATTORNEY/AGENT INFORMATION:  
; NAME: KELLY, ROBERT L.  
; REGISTRATION NUMBER: 31,843  
; REFERENCE/DOCKET NUMBER: 61,686-  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 810-540-0849  
; TELEFAX: 810-540-0763  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 264 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-562-311-4  
Query Match 13.0%; Score 88.5; DB 1; Length 264;  
Best Local Similarity 26.1%; Pred. No. 0.0089;  
Matches 35; Conservative 21; Mismatches 63; Indels 15; Gaps 4;  
QY 5 LEVPCSHALPQGLSPQGVIIVRGLVLOEPKHTVSLR---DQAAHAPVTLRASFA 61  
Db 128 LTVPVDLPLPGGLMPRMLITMGTVKPNANRIVLDFRGNDAFH----FNPRFNNRR 183



Search completed: October 17, 2005, 08:43:38  
Job time : 44 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2005, 08:36:43 ; Search time 168 Seconds

(without alignments)  
329.991 Million cell updates/sec

Title: HJACE54

Perfect score: 682

Sequence: 1 msprlevpcshalpqlspg.....leqlreirsgsvqlcvs 133

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1860064 seqs, 416930855 residues

Total number of hits satisfying chosen parameters: 1860064

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 100 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US11A\_PUBCOMB.pep.\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
- 21: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 22: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	671	98.4	133	14	US-10-156-136-29
2	671	98.4	133	15	US-10-455-366-2
3	671	98.4	133	20	US-11-041-419-29
4	671	98.4	275	15	US-10-455-366-25
5	671	98.4	296	15	US-10-455-366-27
6	671	98.4	336	9	US-09-747-804-1
7	120.5	17.7	196	9	US-09-768-826-55
8	120.5	17.7	196	16	US-10-874-484-55
9	120.5	17.7	318	15	US-10-151-166-4
10	119	17.4	145	9	US-09-728-479-8
11	119	17.4	145	9	US-09-894-526-5
Sequence 29, Appli					
Sequence 2, Appli					
Sequence 29, Appli					
Sequence 25, Appli					
Sequence 27, Appli					
Sequence 1, Appli					
Sequence 55, Appli					
Sequence 55, Appli					
Sequence 4, Appli					
Sequence 8, Appli					
Sequence 5, Appli					

12	119	17.4	145	9	US-09-263-689-12	Sequence 12, Appli
13	119	17.4	145	14	US-10-235-674-12	Sequence 12, Appli
14	119	17.4	145	14	US-10-156-136-42	Sequence 42, Appli
15	119	17.4	145	15	US-10-455-366-3	Sequence 3, Appli
16	119	17.4	145	20	US-11-041-419-42	Sequence 42, Appli
17	109.5	16.1	158	17	US-10-491-213-9	Sequence 9, Appli
18	109.5	16.1	316	16	US-10-856-888-5	Sequence 5, Appli
19	109.5	16.1	316	16	US-10-856-888-15	Sequence 15, Appli
20	109.5	16.1	317	9	US-09-263-689-6	Sequence 6, Appli
21	109.5	16.1	317	14	US-10-235-674-6	Sequence 6, Appli
22	109.5	16.1	318	14	US-10-156-136-43	Sequence 43, Appli
23	109.5	16.1	318	14	US-10-376-133-19	Sequence 19, Appli
24	109.5	16.1	318	15	US-10-170-385-365	Sequence 365, Appli
25	109.5	16.1	318	15	US-10-455-366-4	Sequence 4, Appli
26	109.5	16.1	318	20	US-11-041-419-43	Sequence 43, Appli
27	108.5	15.9	316	9	US-09-747-804-5	Sequence 5, Appli
28	108.5	15.9	316	14	US-10-177-293-354	Sequence 354, Appli
29	108.5	15.9	316	17	US-10-482-029-134	Sequence 134, Appli
30	108.5	15.9	317	10	US-09-948-227-6	Sequence 6, Appli
31	108.5	15.9	317	16	US-10-795-927-6	Sequence 6, Appli
32	108.5	15.9	317	18	US-10-955-119-30	Sequence 30, Appli
33	106.5	15.6	368	16	US-10-856-888-7	Sequence 7, Appli
34	102	15.0	322	9	US-09-728-479-11	Sequence 11, Appli
35	102	15.0	322	18	US-10-958-169-17	Sequence 17, Appli
36	102	15.0	353	15	US-10-633-035-8	Sequence 8, Appli
37	102	15.0	353	18	US-10-958-169-18	Sequence 18, Appli
38	96.5	14.1	358	16	US-10-856-888-6	Sequence 6, Appli
39	96.5	14.1	358	16	US-10-856-888-16	Sequence 16, Appli
40	96.5	14.1	359	14	US-10-376-133-20	Sequence 20, Appli
41	96.5	14.1	359	18	US-10-712-124-92	Sequence 92, Appli
42	95.5	14.0	125	9	US-09-768-826-36	Sequence 36, Appli
43	95.5	14.0	125	11	US-09-833-245-2256	Sequence 256, Appli
44	95.5	14.0	125	16	US-10-874-484-36	Sequence 36, Appli
45	95.5	14.0	135	16	US-10-856-888-20	Sequence 20, Appli
46	95.5	14.0	136	9	US-09-728-479-9	Sequence 9, Appli
47	95.5	14.0	136	9	US-09-154-750A-79	Sequence 79, Appli
48	95.5	14.0	136	9	US-09-263-689-13	Sequence 13, Appli
49	95.5	14.0	136	14	US-10-235-674-13	Sequence 13, Appli
50	95.5	14.0	136	14	US-10-316-253-260	Sequence 260, Appli
51	95.5	14.0	136	15	US-10-133-234A-5	Sequence 2, Appli
52	95.5	14.0	136	15	US-10-398-519-15	Sequence 15, Appli
53	94.5	13.9	131	16	US-09-948-227-7	Sequence 7, Appli
54	94.5	13.9	131	16	US-10-795-927-7	Sequence 7, Appli
55	94	13.8	315	16	US-10-856-888-17	Sequence 17, Appli
56	93	13.6	262	15	US-10-133-234A-5	Sequence 5, Appli
57	91.5	13.4	262	16	US-10-856-888-9	Sequence 9, Appli
58	88.5	13.0	264	16	US-10-473-127-1448	Sequence 1448, Appli
59	87.5	12.8	300	15	US-10-398-519-17	Sequence 15, Appli
60	87.5	12.8	316	9	US-09-263-689-15	Sequence 15, Appli
61	87.5	12.8	316	9	US-09-263-689-17	Sequence 17, Appli
62	87.5	12.8	316	14	US-10-235-674-15	Sequence 15, Appli
63	87.5	12.8	316	14	US-10-235-674-17	Sequence 8, Appli
64	87.5	12.8	316	18	US-10-505-769-8	Sequence 8, Appli
65	87	12.8	140	18	US-10-450-763-53319	Sequence 53319, Appli
66	87	12.8	355	15	US-10-138-588-76	Sequence 76, Appli
67	86.5	12.7	333	17	US-10-506-047-2	Sequence 2, Appli
68	85.5	12.5	263	15	US-10-398-519-14	Sequence 14, Appli
69	84.5	12.4	315	9	US-09-728-479-10	Sequence 10, Appli
70	84.5	12.4	316	15	US-10-151-166-2	Sequence 2, Appli
71	84	12.3	241	9	US-09-738-626-6010	Sequence 6010, Appli
72	84	12.3	324	9	US-09-728-479-7	Sequence 7, Appli
73	84	12.3	324	9	US-09-263-689-11	Sequence 11, Appli
74	84	12.3	324	14	US-10-235-674-11	Sequence 11, Appli
75	82.5	12.1	136	15	US-10-133-234A-7	Sequence 7, Appli
76	82.5	12.1	136	15	US-10-133-234A-8	Sequence 8, Appli
77	82.5	12.1	250	16	US-10-473-127-1350	Sequence 1350, Appli
78	82.5	12.1	281	14	US-10-362-017-5	Sequence 5, Appli
79	81.5	12.0	145	14	US-10-362-017-4	Sequence 4, Appli
80	81.5	12.0	268	14	US-10-106-698-6399	Sequence 2, Appli
81	81.5	12.0	323	9	US-09-802-674-2	Sequence 2, Appli
82	81.5	12.0	323	9	US-09-922-217-1064	Sequence 1064, Appli
83	81.5	12.0	323	9	US-09-988-252-16	Sequence 16, Appli
84	81.5	12.0	323	9	US-09-833-263-1064	Sequence 1064, Appli

85 81.5 12.0 323 9 US-09-263-689-2 Sequence 2, Appli  
86 81.5 12.0 323 9 US-09-981-353-110 Sequence 110, App  
87 81.5 12.0 323 13 US-10-025-380-1064 Sequence 1064, Ap  
88 81.5 12.0 323 14 US-10-235-674-2 Sequence 2, Appli  
89 81.5 12.0 323 14 US-10-235-994-30 Sequence 30, Appl  
90 81.5 12.0 323 14 US-10-376-133-17 Sequence 17, Appl  
91 81.5 12.0 323 16 US-10-776-601-16 Sequence 16, Appl  
92 81.5 12.0 323 16 US-10-856-888-21 Sequence 21, Appl  
93 81.5 12.0 323 20 US-11-032-421-2 Sequence 2, Appli  
94 81.5 12.0 329 9 US-09-802-674-13 Sequence 13, Appl  
95 81.5 12.0 329 20 US-11-032-421-13 Sequence 13, Appl  
96 81.5 12.0 351 14 US-10-106-698-4502 Sequence 4502, Ap  
97 81 11.9 171 14 US-10-190-866A-2 Sequence 2, Appli  
98 81 11.9 401 14 US-10-190-866A-1 Sequence 1, Appli  
99 80.5 11.8 242 15 US-10-133-234A-4 Sequence 4, Appli  
100 80 11.7 145 9 US-09-894-526-1 Sequence 1, Appli

ALIGNMENTS

RESULT 1  
US-10-156-136-29  
; Sequence 29, Application US/10156136  
; Publication No. US20030129696A1  
; GENERAL INFORMATION:  
; APPLICANT: Ni, et al.  
; TITLE OF INVENTION: Polynucleotides and Polypeptides Encoding Receptors  
; NUMBER OF SEQUENCES: 63  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Ave  
; CITY: Rockville  
; STATE: MD  
; COUNTRY: US  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/156,136  
; FILING DATE: 29-May-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/09/010,146  
; FILING DATE: <Unknown>  
; APPLICATION NUMBER: WO US98/00959  
; FILING DATE: 21-JAN-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: A. Anders Brookes  
; REGISTRATION NUMBER: 36,373  
; REFERENCE/DOCKET NUMBER: PF354PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 301-309-8504  
; TELEFAX: 301-309-8439  
; INFORMATION FOR SEQ ID NO: 29:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 133 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 29:  
US-10-156-136-29  
Query Match 98.4%; Score 671; DB 14; Length 133;  
Best Local Similarity 99.2%; Pred. No. 6.1e-68;  
Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
1 MSPRLEVPCHALPQGLSPGVIIVRGLVLPQPKHFTVSLRDQAAHAPVTLRASFDRTL 60

Db 1 MSPRLEVPCHALPQGLSPGVIIVRGLVLPQPKHFTVSLRDQAAHAPVTLRASFDRTL 60  
QY 61 AMISRWGQKLLISAPPLFYPPQPFVLLLFQSGGLKIALNGOGLGATSMNQALEQLREL 120  
Db 61 AMISRWGQKLLISAPPLFYPPQPFVLLLFQSGGLKIALNGOGLGATSMNQALEQLREL 120  
QY 121 RISGSVQLYCVHS 133  
Db 121 RISGSVQLYCVHS 133  
RESULT 2  
US-10-455-366-2  
; Sequence 2, Application US/10455366  
; Publication No. US20030208044A1  
; GENERAL INFORMATION:  
; APPLICANT: Ni et al.  
; TITLE OF INVENTION: Galectin 11  
; FILE REFERENCE: PF354P2  
; CURRENT APPLICATION NUMBER: US/10/455,366  
; CURRENT FILING DATE: 2003-06-06  
; PRIOR APPLICATION NUMBER: US/09/557,170  
; PRIOR FILING DATE: 2000-04-21  
; PRIOR APPLICATION NUMBER: 09/109,864  
; PRIOR FILING DATE: 1998-06-06  
; PRIOR APPLICATION NUMBER: 09/010,146  
; PRIOR FILING DATE: 1998-01-21  
; PRIOR APPLICATION NUMBER: 60/034,205  
; PRIOR FILING DATE: 1997-01-21  
; PRIOR APPLICATION NUMBER: 60/034,204  
; PRIOR FILING DATE: 1997-01-21  
; PRIOR APPLICATION NUMBER: 60/169,932  
; PRIOR FILING DATE: 1999-12-10  
; PRIOR APPLICATION NUMBER: 60/130,390  
; PRIOR FILING DATE: 1999-04-21  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 133  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-455-366-2  
Query Match 98.4%; Score 671; DB 15; Length 133;  
Best Local Similarity 99.2%; Pred. No. 6.1e-68;  
Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
1 MSPRLEVPCHALPQGLSPGVIIVRGLVLPQPKHFTVSLRDQAAHAPVTLRASFDRTL 60  
Db 1 MSPRLEVPCHALPQGLSPGVIIVRGLVLPQPKHFTVSLRDQAAHAPVTLRASFDRTL 60  
QY 61 AMISRWGQKLLISAPPLFYPPQPFVLLLFQSGGLKIALNGOGLGATSMNQALEQLREL 120  
Db 61 AMISRWGQKLLISAPPLFYPPQPFVLLLFQSGGLKIALNGOGLGATSMNQALEQLREL 120  
QY 121 RISGSVQLYCVHS 133  
Db 121 RISGSVQLYCVHS 133  
RESULT 3  
US-11-041-419-29  
; Sequence 29, Application US/11041419  
; Publication No. US20050176044A1  
; GENERAL INFORMATION:  
; APPLICANT: Ni, et al.  
; TITLE OF INVENTION: Polynucleotides And Polypeptides Encoding Receptors  
; FILE REFERENCE: PF354C3  
; CURRENT APPLICATION NUMBER: US/11/041,419  
; CURRENT FILING DATE: 2005-01-25  
; NUMBER OF SEQ ID NOS: 63  
; SEQ ID NO 29

```
; LENGTH: 133
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-041-419-29

Query Match      98.4%; Score 671; DB 20; Length 133;
Best Local Similarity 99.2%; Pred. No. 6.1e-68;
Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 MSPRLEVPCHALPQGLSPGOVIIVRGLVLOEPKHPTVSLRDQAAHAPVTLRASFAADRTL 60
        |||||||
Db      1 MSPRLEVPCHALPQGLSPGOVIIVRGLVLOEPKHPTVSLRDQAAHAPVTLRASFAADRTL 60
        |||||||

Qy      61 AWISRWGQKKLISAPFLFYQRFPEVLLFQEGGKLALNGQGLGATSMNQALEQREL 120
        |||||||
Db      61 AWISRWGQKKLISAPFLFYQRFPEVLLFQEGGKLALNGQGLGATSMNQALEQREL 120
        |||||||

Qy      121 RISGSVQLYCVHS 133
        |||||||
Db      121 RISGSVQLYCVHS 133
        |||||||

RESULT 4
US-10-455-366-25
; Sequence 25, Application US/10455366
; Publication No. US20030208044A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: Galectin 11
; FILE REFERENCE: PF354P2
; CURRENT APPLICATION NUMBER: US/10/455,366
; CURRENT FILING DATE: 2003-06-06
; PRIOR APPLICATION NUMBER: US/09/557,170
; PRIOR FILING DATE: 2000-04-21
; PRIOR APPLICATION NUMBER: 09/109,864
; PRIOR FILING DATE: 1998-06-06
; PRIOR APPLICATION NUMBER: 60/034,205
; PRIOR FILING DATE: 1997-01-21
; PRIOR APPLICATION NUMBER: 60/034,204
; PRIOR FILING DATE: 1997-01-21
; PRIOR APPLICATION NUMBER: 60/169,932
; PRIOR FILING DATE: 1999-12-10
; PRIOR APPLICATION NUMBER: 60/130,390
; PRIOR FILING DATE: 1999-04-21
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 25
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-455-366-25

Query Match      98.4%; Score 671; DB 15; Length 275;
Best Local Similarity 99.2%; Pred. No. 1.5e-67;
Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 MSPRLEVPCHALPQGLSPGOVIIVRGLVLOEPKHPTVSLRDQAAHAPVTLRASFAADRTL 60
        |||||||
Db      143 MSPRLEVPCHALPQGLSPGOVIIVRGLVLOEPKHPTVSLRDQAAHAPVTLRASFAADRTL 202
        |||||||

Qy      61 AWISRWGQKKLISAPFLFYQRFPEVLLFQEGGKLALNGQGLGATSMNQALEQREL 120
        |||||||
Db      203 AWISRWGQKKLISAPFLFYQRFPEVLLFQEGGKLALNGQGLGATSMNQALEQREL 262
        |||||||

Qy      121 RISGSVQLYCVHS 133
        |||||||
Db      263 RISGSVQLYCVHS 275
        |||||||

RESULT 5
US-10-455-366-27
```

```
; Sequence 27, Application US/10455366
; Publication No. US20030208044A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: Galectin 11
; FILE REFERENCE: PF354P2
; CURRENT APPLICATION NUMBER: US/10/455,366
; CURRENT FILING DATE: 2003-06-06
; PRIOR APPLICATION NUMBER: US/09/557,170
; PRIOR FILING DATE: 2000-04-21
; PRIOR APPLICATION NUMBER: 09/109,864
; PRIOR FILING DATE: 1998-06-06
; PRIOR APPLICATION NUMBER: 09/010,146
; PRIOR FILING DATE: 1998-01-21
; PRIOR APPLICATION NUMBER: 60/034,205
; PRIOR FILING DATE: 1997-01-21
; PRIOR APPLICATION NUMBER: 60/034,204
; PRIOR FILING DATE: 1997-01-21
; PRIOR APPLICATION NUMBER: 60/169,932
; PRIOR FILING DATE: 1999-12-10
; PRIOR APPLICATION NUMBER: 60/130,390
; PRIOR FILING DATE: 1999-04-21
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 27
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-455-366-27

Query Match      98.4%; Score 671; DB 15; Length 296;
Best Local Similarity 99.2%; Pred. No. 1.7e-67;
Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 MSPRLEVPCHALPQGLSPGOVIIVRGLVLOEPKHPTVSLRDQAAHAPVTLRASFAADRTL 60
        |||||||
Db      164 MSPRLEVPCHALPQGLSPGOVIIVRGLVLOEPKHPTVSLRDQAAHAPVTLRASFAADRTL 223
        |||||||

Qy      61 AWISRWGQKKLISAPFLFYQRFPEVLLFQEGGKLALNGQGLGATSMNQALEQREL 120
        |||||||
Db      224 AWISRWGQKKLISAPFLFYQRFPEVLLFQEGGKLALNGQGLGATSMNQALEQREL 283
        |||||||

Qy      121 RISGSVQLYCVHS 133
        |||||||
Db      284 RISGSVQLYCVHS 296
        |||||||

RESULT 6
US-09-747-804-1
; Sequence 1, Application US/09747804
; Patent No. US20010010913A1
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Patterson, Chandra
; TITLE OF INVENTION: EXTRACELLULAR ADHESIVE PROTEINS
; FILE REFERENCE: PF-0576 US
; CURRENT APPLICATION NUMBER: US/09/747,804
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 09/131,648
; PRIOR FILING DATE: 1998-08-10
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: 2635136
US-09-747-804-1
```







```
Matches 35; Conservative 26; Mismatches 65; Indels 4; Gaps 2;
QY 3 PRLEVPCHALPQGLSPGVIIVRGLVLPQPKPTVSLR---DQAAHAPVTLRASPADRT 59
Db 11 PNLAVPFTSIPNGLYPSKSIIVSGVLSDAKRFQINLRCCGDIAPHLNPRFDENAVVRN 70
QY 60 LAWISRWG-QKKLISAPFLFYPPORFVLLLFQEGGLKALNGGGLGATSMNQQALEQLR 118
Db 71 TQINNSWGPERSLPGSMPPSRGQRFVSWILCEGHCFKVAVDGQHICEYSHRLMNLDPIN 130
QY 119 ELRISGVQL 128
Db 131 TLEVAGDIQL 140

RESULT 13
US-10-235-674-12
; Sequence 12, Application US/10235674
; Publication No. US20030040081A1
; GENERAL INFORMATION:
; APPLICANT: Ni, Jian
; APPLICANT: Gentz, Reiner L.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Galectin 9 and 10SV Polynucleotides
; FILE REFERENCE: 1488.0560004
; CURRENT APPLICATION NUMBER: US/10/235,674
; CURRENT FILING DATE: 2002-09-06
; PRIOR APPLICATION NUMBER: US 09/656,450
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: US 09/263,689
; PRIOR FILING DATE: 1999-03-05
; PRIOR APPLICATION NUMBER: US 08/946,914
; PRIOR FILING DATE: 1997-10-09
; PRIOR APPLICATION NUMBER: US 60/028,093
; PRIOR FILING DATE: 1996-10-09
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Rat
US-10-235-674-12

Query Match 17.4%; Score 119; DB 14; Length 145;
Best Local Similarity 26.9%; Pred. No. 3.6e-05;
Matches 35; Conservative 26; Mismatches 65; Indels 4; Gaps 2;
QY 3 PRLEVPCHALPQGLSPGVIIVRGLVLPQPKPTVSLR---DQAAHAPVTLRASPADRT 59
Db 11 PNLAVPFTSIPNGLYPSKSIIVSGVLSDAKRFQINLRCCGDIAPHLNPRFDENAVVRN 70
QY 60 LAWISRWG-QKKLISAPFLFYPPORFVLLLFQEGGLKALNGGGLGATSMNQQALEQLR 118
Db 71 TQINNSWGPERSLPGSMPPSRGQRFVSWILCEGHCFKVAVDGQHICEYSHRLMNLDPIN 130
QY 119 ELRISGVQL 128
Db 131 TLEVAGDIQL 140

RESULT 14
US-10-156-136-42
; Sequence 42, Application US/10156136
; Publication No. US20030129696A1
; GENERAL INFORMATION:
; APPLICANT: Ni, et al.
; TITLE OF INVENTION: Polynucleotides and Polypeptides Encoding Receptors
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Ave
; CITY: Rockville
```

```
; STATE: MD
; COUNTRY: US
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/10/156,136
; APPLICATION NUMBER: US/10/156,136
; FILING DATE: 29-May-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/010,146
; FILING DATE: <Unknown>
; APPLICATION NUMBER: WO US98/00959
; FILING DATE: 21-JAN-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brookes
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PF354PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-309-8504
; TELEFAX: 301-309-8439
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 145 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 42:
US-10-156-136-42

Query Match 17.4%; Score 119; DB 14; Length 145;
Best Local Similarity 26.9%; Pred. No. 3.6e-05;
Matches 35; Conservative 26; Mismatches 65; Indels 4; Gaps 2;
QY 3 PRLEVPCHALPQGLSPGVIIVRGLVLPQPKPTVSLR---DQAAHAPVTLRASPADRT 59
Db 11 PNLAVPFTSIPNGLYPSKSIIVSGVLSDAKRFQINLRCCGDIAPHLNPRFDENAVVRN 70
QY 60 LAWISRWG-QKKLISAPFLFYPPORFVLLLFQEGGLKALNGGGLGATSMNQQALEQLR 118
Db 71 TQINNSWGPERSLPGSMPPSRGQRFVSWILCEGHCFKVAVDGQHICEYSHRLMNLDPIN 130
QY 119 ELRISGVQL 128
Db 131 TLEVAGDIQL 140

RESULT 15
US-10-455-366-3
; Sequence 3, Application US/10455366
; Publication No. US20030208044A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: Galectin 11
; FILE REFERENCE: PF354P2
; CURRENT APPLICATION NUMBER: US/10/455,366
; CURRENT FILING DATE: 2003-06-06
; PRIOR APPLICATION NUMBER: US/09/557,170
; PRIOR FILING DATE: 2000-04-21
; PRIOR APPLICATION NUMBER: 09/109,864
; PRIOR FILING DATE: 1998-06-06
; PRIOR APPLICATION NUMBER: 09/010,146
; PRIOR FILING DATE: 1998-01-21
; PRIOR APPLICATION NUMBER: 60/034,205
; PRIOR FILING DATE: 1997-01-21
; PRIOR APPLICATION NUMBER: 60/034,204
; PRIOR FILING DATE: 1997-01-21
; PRIOR APPLICATION NUMBER: 60/169,932
; PRIOR FILING DATE: 1999-12-10
```

; PRIOR APPLICATION NUMBER: 60/130,390
; PRIOR FILING DATE: 1999-04-21
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-455-366-3

Query Match 17.4%; Score 119; DB 15; Length 145;
Best Local Similarity 26.9%; Pred. No. 3.6e-05;
Matches 35; Conservative 26; Mismatches 65; Indels 4; Gaps 2;

Qy 3 PRLEVPCHALPQGLSPGVIIIVRGLVLOEPKHTVSLR---DOAAHAPVTLRASPADRT 59
Db 11 PNLAVPFTSIPNGLYPSKISIVISGVLSDAKRFQINLRGGDIAPHLNPRFDENAVVRN 70

Qy 60 LAWISRWG-QKKLISAPFLFPQRFVLLLFQEGGKLALNGGGLGATSMNQQAELQLR 118
Db 71 TQINNSWGPEERSLPGSPFSGRQRFVSWILCEGHCFKVAVDGQHICEYSHRLMLNLPDIN 130

Qy 119 ELRISGSVOL 128
Db 131 TLEVAGDIQL 140

RESULT 16
US-11-041-419-42
; Sequence 42, Application US/11041419
; Publication No. US20050176044A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: Polynucleotides And Polypeptides Encoding Receptors
; FILE REFERENCE: PF354C3
; CURRENT APPLICATION NUMBER: US/11/041.419
; CURRENT FILING DATE: 2005-01-25
; NUMBER OF SEQ ID NOS: 63
; SEQ ID NO 42
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Rattus sp.
US-11-041-419-42

Query Match 17.4%; Score 119; DB 20; Length 145;
Best Local Similarity 26.9%; Pred. No. 3.6e-05;
Matches 35; Conservative 26; Mismatches 65; Indels 4; Gaps 2;

Qy 3 PRLEVPCHALPQGLSPGVIIIVRGLVLOEPKHTVSLR---DOAAHAPVTLRASPADRT 59
Db 11 PNLAVPFTSIPNGLYPSKISIVISGVLSDAKRFQINLRGGDIAPHLNPRFDENAVVRN 70

Qy 60 LAWISRWG-QKKLISAPFLFPQRFVLLLFQEGGKLALNGGGLGATSMNQQAELQLR 118
Db 71 TQINNSWGPEERSLPGSPFSGRQRFVSWILCEGHCFKVAVDGQHICEYSHRLMLNLPDIN 130

Qy 119 ELRISGSVOL 128
Db 131 TLEVAGDIQL 140

RESULT 17
US-10-491-213-9
; Sequence 9, Application US/10491213
; Publication No. US20050048490A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION; AZIMZAI, Yalda;
; APPLICANT: BAUGHN, Mariah R.; BECHA, Shanya D.;
; APPLICANT: BOWORSKY, Mark L.; CHAWLA, Narinder K.;
; APPLICANT: ELLIOTT, Vicki S.; EMERLING, Brooke M.;
; APPLICANT: GANDHI, Ameena R.; GIETZEN, Kimberly J.;
; APPLICANT: GORVAD, Ann E.; GRIFFIN, Jennifer A.;
; APPLICANT: HAFALIA, April J.A.; ISON, Craig H.;

; APPLICANT: KABLE, Amy E.; KALAFUS, Daniel P.;
; APPLICANT: LEHR-MASON, Patricia M.; LU, Dyung Aina M.;
; APPLICANT: MARQUIS, Joseph P.; NGUYEN, Daniel B.;
; APPLICANT: RAMKUMAR, Jeyalaksmi; RICHARDSON, Thomas W.;
; APPLICANT: KAREHT, Stephanie K.; SWARNAKAR, Anita;
; APPLICANT: TANG, Y. Tom; TRAN, Uyen K.;
; APPLICANT: WARREN, Bridget A.; XU, Yuming;
; APPLICANT: YAO, Monique G.; YUE, Huibin;
; APPLICANT: YUE, Henry
; TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL GROWTH, DIFFERENTIATION, AND DEATH
; FILE REFERENCE: PF-1213 USN
; CURRENT APPLICATION NUMBER: US/10/491,213
; CURRENT FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: PCT/US02/31095
; PRIOR FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/326,389
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US 60/327,380
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: US 60/328,186
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: US 60/329,690
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/345,384
; PRIOR FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: US 60/348,165
; PRIOR FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: US 60/350,219
; PRIOR FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: US 60/344,518
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: US 60/345,143
; PRIOR FILING DATE: 2001-11-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PERL Program
; SEQ ID NO 9
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: incyte ID No: 7500513CDI
US-10-491-213-9

Query Match 16.1%; Score 109.5; DB 17; Length 158;
Best Local Similarity 25.5%; Pred. No. 0.00048;
Matches 35; Conservative 23; Mismatches 74; Indels 5; Gaps 1;

Qy 2 SPRLEVPCHALPQGLSPGVIIIVRGLVLOEPKHTVSL-----RDOAAHAPVTLRASFA 56
Db 21 TPQLSLFFAARLNTPMGPGRTVVVKGVNANAKSFNVDLLAGSKDIALHLNRLNIKAF 80

Qy 57 DRTLAWISRWGQKKLISAPFLFPQRFVLLLFQEGGKLALNGGGLGATSMNQQAELQ 116
Db 81 VRNSFLOESGSEERNTTSFPFSGMYFEMIIICYDREFKVAVNGVHSLEYKHFRLKLS 140

Qy 117 LRELISGSVOLYCVHS 133
Db 141 IDTLEINGDIHLEVR 157

RESULT 18
US-10-856-888-5
; Sequence 5, Application US/10856888
; Publication No. US20040235037A1
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; TITLE OF INVENTION: Polymorphic Markers of Prostate Carcinoma Tumor Antigen-1 (PCTA-1)
; FILE REFERENCE: GEN-1112XC1
; CURRENT APPLICATION NUMBER: US/10/856,888

```

; CURRENT FILING DATE: 2004-05-27
; PRIOR APPLICATION NUMBER: US/09/326,402
; PRIOR FILING DATE: 1999-06-04
; PRIOR APPLICATION NUMBER: 60/088,187
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/102,324
; PRIOR FILING DATE: 1998-09-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 16
; OTHER INFORMATION: 5-2-162 : polymorphic amino acid Tyr or Phe
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 35
; OTHER INFORMATION: 5-2-213 : polymorphic amino acid Cys or Arg
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 55
; OTHER INFORMATION: 5-3-84 : polymorphic amino acid Val or Met
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 163
; OTHER INFORMATION: 5-7-195 : polymorphic amino acid Ser or Arg
; US-10-856-888-5

Query Match      16.1%; Score 109.5; DB 16; Length 316;
Best Local Similarity 25.5%; Pred. No. 0.0011;
Matches 35; Conservative 23; Mismatches 74; Indels 5; Gaps 1;

QY      2 SPRLEVPCHALPQGLSPGQVILVRGLVLPQEPKHTVSL-----RDQAAHAPVTLRASFA 56
Db      179 TPQLSLPFAARLNTPMGPGRTVVVKGVEVNAKSFNVDDLAKGSKDIALHLNPLRIKAP 238
QY      57 DRTLAWISWGQKKLISAPFLFPQRFVFLVLLFQGGKLALNGQGLGATSMNQQALEQ 116
Db      239 VRNSFLOESWGEEERNITSPFSPGMYFEMIIYCDVREFKVAVNGVHSLKYKHFKELESS 298
QY      117 LRELRISSGVOLYCVHS 133
Db      299 IDTLEINGDIHLLVRS 315

RESULT 19
US-10-856-888-15
; Sequence 15, Application US/10856888
; Publication No. US20040235037A1
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; TITLE OF INVENTION: Polymorphic Markers of Prostate Carcinoma Tumor Antigen-1 (PCTA-1
; FILE REFERENCE: GEN-FLIXC1
; CURRENT APPLICATION NUMBER: US/10/856,888
; PRIOR FILING DATE: 2004-05-27
; PRIOR APPLICATION NUMBER: US/09/326,402
; PRIOR FILING DATE: 1999-06-04
; PRIOR APPLICATION NUMBER: 60/088,187
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/102,324
; PRIOR FILING DATE: 1998-09-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:

Query Match      16.1%; Score 109.5; DB 9; Length 317;
Best Local Similarity 25.5%; Pred. No. 0.0011;
Matches 35; Conservative 23; Mismatches 74; Indels 5; Gaps 1;

QY      2 SPRLEVPCHALPQGLSPGQVILVRGLVLPQEPKHTVSL-----RDQAAHAPVTLRASFA 56
Db      180 TPQLSLPFAARLNTPMGPGRTVVVKGVEVNAKSFNVDDLAKGSKDIALHLNPLRIKAP 239
; FEATURE:

; NAME/KEY: MISC FEATURE
; LOCATION: (1)-(316)
; OTHER INFORMATION: amino acid sequence of PCTA
; US-10-856-888-15

Query Match      16.1%; Score 109.5; DB 16; Length 316;
Best Local Similarity 25.5%; Pred. No. 0.0011;
Matches 35; Conservative 23; Mismatches 74; Indels 5; Gaps 1;

QY      2 SPRLEVPCHALPQGLSPGQVILVRGLVLPQEPKHTVSL-----RDQAAHAPVTLRASFA 56
Db      179 TPQLSLPFAARLNTPMGPGRTVVVKGVEVNAKSFNVDDLAKGSKDIALHLNPLRIKAP 238
QY      57 DRTLAWISWGQKKLISAPFLFPQRFVFLVLLFQGGKLALNGQGLGATSMNQQALEQ 116
Db      239 VRNSFLOESWGEEERNITSPFSPGMYFEMIIYCDVREFKVAVNGVHSLKYKHFKELESS 298
QY      117 LRELRISSGVOLYCVHS 133
Db      299 IDTLEINGDIHLLVRS 315

RESULT 20
US-09-263-689-6
; Sequence 6, Application US/09263689
; Patent No. US20020150970A1
; GENERAL INFORMATION:
; APPLICANT: Ni, Jian
; APPLICANT: Gentz, Reiner L.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Galectin 8, 9, 10 and 10SV
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein, & Fox P.L.L.C.
; STREET: 1100 New York Ave., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/263,689
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/946,914
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Steffe, Eric K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.0560001/EKS/SGW
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 317 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-263-689-6

Query Match      16.1%; Score 109.5; DB 9; Length 317;
Best Local Similarity 25.5%; Pred. No. 0.0011;
Matches 35; Conservative 23; Mismatches 74; Indels 5; Gaps 1;

QY      2 SPRLEVPCHALPQGLSPGQVILVRGLVLPQEPKHTVSL-----RDQAAHAPVTLRASFA 56
Db      180 TPQLSLPFAARLNTPMGPGRTVVVKGVEVNAKSFNVDDLAKGSKDIALHLNPLRIKAP 239
; FEATURE:
```

```
Qy 57 DRTLAWISRWGQKLLISAPFLFYQRFVLLFQEGGLKALNGOGLGATSMNQALEQ 116
      ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||:::
Db 240 VRNSFLQESWGEEERNITAPFSPGMYFEMIIYCDVREFKVAVNGVSHSLEYKHFKESS 299
      : ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||:::
Qy 117 LRELRISSGVOLYCVHS 133
      : ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||:::
Db 300 IDTLEINGDIHLLVRS 316
      : ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||:::

RESULT 21
US-10-235-674-6
; Sequence 6, Application US/10235674
; Publication No. US20030040081A1
; GENERAL INFORMATION:
; APPLICANT: Ni, Jian
; APPLICANT: Gentz, Reiner L.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Galectin 9 and 10SV Polynucleotides
; FILE REFERENCE: 1488.0560004
; CURRENT APPLICATION NUMBER: US/10/235,674
; CURRENT FILING DATE: 2002-09-06
; PRIOR APPLICATION NUMBER: US 09/656,450
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: US 09/263,689
; PRIOR FILING DATE: 1999-03-05
; PRIOR APPLICATION NUMBER: US 08/946,914
; PRIOR FILING DATE: 1997-10-09
; PRIOR APPLICATION NUMBER: US 60/028,093
; PRIOR FILING DATE: 1996-10-09
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 317
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-235-674-6

Query Match 16.1%; Score 109.5; DB 14; Length 317;
Best Local Similarity 25.5%; Pred. No. 0.0011;
Matches 35; Conservative 23; Mismatches 74; Indels 5; Gaps 1;

Qy 2 SPRLEVPCHALPQGLSPGQVIVRGLVLOBPKHPTVSL-----RDOAAHAPVTLRASFA 56
      ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||:::
Db 180 TPQLSLPFAARLNTPMGRTVVVKGEVNAKSFNVDLLAGKSKDIALHLPRLNIKAF 239
      : ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||:::

Qy 57 DRTLAWISRWGQKLLISAPFLFYQRFVLLFQEGGLKALNGOGLGATSMNQALEQ 116
      ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||:::
Db 240 VRNSFLQESWGEEERNITAPFSPGMYFEMIIYCDVREFKVAVNGVSHSLEYKHFKESS 299
      : ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||:::

Qy 117 LRELRISSGVOLYCVHS 133
      : ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||:::
Db 300 IDTLEINGDIHLLVRS 316
      : ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||:::

RESULT 22
US-10-156-136-43
; Sequence 43, Application US/10156136
; Publication No. US20030129696A1
; GENERAL INFORMATION:
; APPLICANT: Ni, et al.
; TITLE OF INVENTION: Polynucleotides and Polypeptides Encoding
; CORRESPONDENCE ADDRESSES: 63
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Ave
; CITY: Rockville
; STATE: MD
; COUNTRY: US
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/156.136
FILING DATE: 29-May-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/010.146
FILING DATE: <Unknown>
APPLICATION NUMBER: WO US98/00959
FILING DATE: 21-JAN-1998
ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brookes
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PF354PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-309-8504
TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 318 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 43:
US-10-156-136-43

Query Match 16.1%; Score 109.5; DB 14; Length 318;
Best Local Similarity 25.5%; Pred. No. 0.0011;
Matches 35; Conservative 23; Mismatches 74; Indels 5; Gaps 1;

Qy 2 SPRLEVPCHALPQGLSPGQVIVRGLVLOBPKHPTVSL-----RDOAAHAPVTLRASFA 56
      ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||:::
Db 181 TPQLSLPFAARLNTPMGRTVVVKGEVNAKSFNVDLLAGKSKDIALHLPRLNIKAF 240
      : ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||:::

Qy 57 DRTLAWISRWGQKLLISAPFLFYQRFVLLFQEGGLKALNGOGLGATSMNQALEQ 116
      ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||:::
Db 241 VRNSFLQESWGEEERNITAPFSPGMYFEMIIYCDVREFKVAVNGVSHSLEYKHFKESS 300
      : ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||:::

Qy 117 LRELRISSGVOLYCVHS 133
      : ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||:::
Db 301 IDTLEINGDIHLLVRS 317
      : ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||:::

RESULT 23
US-10-376-133-19
; Sequence 19, Application US/10376133
; Publication No. US20030165965A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: LGALS AS MODIFIERS OF THE CHK PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX03-014C
; CURRENT APPLICATION NUMBER: US/10/376,133
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,757
; PRIOR FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 19
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-376-133-19

Query Match 16.1%; Score 109.5; DB 14; Length 318;
Best Local Similarity 25.5%; Pred. No. 0.0011;
Matches 35; Conservative 23; Mismatches 74; Indels 5; Gaps 1;

Qy 2 SPRLEVPCHALPQGLSPGQVIVRGLVLOBPKHPTVSL-----RDOAAHAPVTLRASFA 56
      ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||:::
Db 181 TPQLSLPFAARLNTPMGRTVVVKGEVNAKSFNVDLLAGKSKDIALHLPRLNIKAF 240
      : ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||:::

Qy 57 DRTLAWISRWGQKLLISAPFLFYQRFVLLFQEGGLKALNGOGLGATSMNQALEQ 116
      ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||:::
Db 241 VRNSFLQESWGEEERNITAPFSPGMYFEMIIYCDVREFKVAVNGVSHSLEYKHFKESS 300
      : ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||:::

Qy 117 LRELRISSGVOLYCVHS 133
      : ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||:::
Db 301 IDTLEINGDIHLLVRS 317
      : ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||:::

RESULT 24
US-10-156-136-43
; Sequence 43, Application US/10156136
; Publication No. US20030129696A1
; GENERAL INFORMATION:
; APPLICANT: Ni, et al.
; TITLE OF INVENTION: Polynucleotides and Polypeptides Encoding
; CORRESPONDENCE ADDRESSES: 63
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Ave
; CITY: Rockville
; STATE: MD
; COUNTRY: US
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```



```

; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Patterson, Chandra
; TITLE OF INVENTION: EXTRACELLULAR ADHESIVE PROTEINS
; FILE REFERENCE: PP-0576 US
; CURRENT APPLICATION NUMBER: US/09/747,804
; PRIOR FILING DATE: 2000-12-22
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 09/131,648
; PRIOR FILING DATE: 1998-08-10
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PERL Program
; SEQ ID NO 5
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: g1932712
US-09-747-804-5

Query Match      15.9%; Score 108.5; DB 9; Length 316;
Best Local Similarity 25.5%; Pred. No. 0.0015;
Matches 35; Conservative 23; Mismatches 74; Indels 5; Gaps 1;

Qy      2  SPRLEVPCHALPQGLSPGQVIVIRGLVLOEPKHPTVSL-----RDQAAHAPVTLRASFA 56
Db      179  TPQLSLPFAARLNTPMGRTVVVQGEVNAKSFNVDDLAKGSKDIALHLNPRLNKAF 238

Qy      57  DRTLAWISRWGQKKLISAPFLFYQRFVEVLLLFQEGGLKALNGOGLGATSMNQALEQ 116
Db      239  VRNSFLOESGGEERNITSPFPSPGMYFEMIYCDVREFKVAVNGVHSLEYKHRFKELSS 298

Qy      117  LRELRISSVOLYCVHS 133
Db      299  IDTLEINGDIHLLVRS 315

RESULT 28
US-10-177-293-354
; Sequence 354, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Gannavarpu, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Pusztai, Lajos
; APPLICANT: Meric, Funda
; APPLICANT: Mills, Gordon B.
; APPLICANT: Sahin, Aysegul
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
; TITLE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10/177,293
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002

```

```

; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 354
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-293-354

Query Match      15.9%; Score 108.5; DB 14; Length 316;
Best Local Similarity 25.5%; Pred. No. 0.0015;
Matches 35; Conservative 23; Mismatches 74; Indels 5; Gaps 1;

Qy      2  SPRLEVPCHALPQGLSPGQVIVIRGLVLOEPKHPTVSL-----RDQAAHAPVTLRASFA 56
Db      179  TPQLSLPFAARLNTPMGRTVVVQGEVNAKSFNVDDLAKGSKDIALHLNPRLNKAF 238

Qy      57  DRTLAWISRWGQKKLISAPFLFYQRFVEVLLLFQEGGLKALNGOGLGATSMNQALEQ 116
Db      239  VRNSFLOESGGEERNITSPFPSPGMYFEMIYCDVREFKVAVNGVHSLEYKHRFKELSS 298

Qy      117  LRELRISSVOLYCVHS 133
Db      299  IDTLEINGDIHLLVRS 315

RESULT 29
US-10-482-029-134
; Sequence 134, Application US/10482029
; Publication No. US20050037445A1
; GENERAL INFORMATION:
; APPLICANT: ODIN medical A/S
; TITLE OF INVENTION: Oncology drug innovation
; FILE REFERENCE: P 573 PC00
; CURRENT APPLICATION NUMBER: US/10/482,029
; CURRENT FILING DATE: 2003-12-29
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 134
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-482-029-134

Query Match      15.9%; Score 108.5; DB 17; Length 316;
Best Local Similarity 25.5%; Pred. No. 0.0015;
Matches 35; Conservative 23; Mismatches 74; Indels 5; Gaps 1;

Qy      2  SPRLEVPCHALPQGLSPGQVIVIRGLVLOEPKHPTVSL-----RDQAAHAPVTLRASFA 56
Db      179  TPQLSLPFAARLNTPMGRTVVVQGEVNAKSFNVDDLAKGSKDIALHLNPRLNKAF 238

Qy      57  DRTLAWISRWGQKKLISAPFLFYQRFVEVLLLFQEGGLKALNGOGLGATSMNQALEQ 116
Db      239  VRNSFLOESGGEERNITSPFPSPGMYFEMIYCDVREFKVAVNGVHSLEYKHRFKELSS 298

Qy      117  LRELRISSVOLYCVHS 133
Db      299  IDTLEINGDIHLLVRS 315

RESULT 30
US-09-948-227-6
; Sequence 6, Application US/09948227
; Publication No. US20030050266A1
; GENERAL INFORMATION:
; APPLICANT: Paul B. Fisher
; APPLICANT: Raul V. Gopalakrishnan
; TITLE OF INVENTION: Anti-tumor Effects of Prostate Carcinoma

```





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2005, 08:32:58 ; Search time 16 Seconds  
(without alignments)  
799.802 Million cell updates/sec

Title: HJACE54  
Perfect score: 682  
Sequence: 1 msprlvpcchalpqlspg.....leqlrelrsgsvqlcvs 133

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : PIR 79:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	119	17.4	145	2	galectin-5 - rat
2	108.5	15.9	317	2	prostate carcinoma
3	95.5	14.0	136	2	galectin-7 - human
4	94.5	13.9	727	2	heterocyst-specifi
5	94.5	13.9	727	2	heterocyst-specifi
6	87.5	12.8	316	2	galectin-8 - rat
7	85.5	12.5	264	2	galactose-specific
8	85.5	12.5	264	2	galactose-specific
9	84	12.3	324	2	lactose-binding le
10	81	11.9	848	2	interleukin-4-indu
11	80.5	11.8	242	2	galectin-3 - rabbi
12	78.5	11.5	262	2	IGF-binding protei
13	78.5	11.5	323	2	lectin L-36 - pig
14	77.5	11.4	245	2	carbohydrate-bind
15	77.5	11.4	250	2	galectin 3 - human
16	76.5	11.2	294	2	lactose-binding le
17	74.5	10.9	129	2	ig kappa chain - h
18	72.5	10.6	191	2	hypothetical prote
19	72.5	10.6	279	2	beta-galactoside-b
20	72.5	10.6	285	2	hypothetical prote
21	72.5	10.6	316	1	nonstructural prot
22	72.5	10.6	645	2	asparagine synthas
23	72	10.6	131	2	T cell receptor Ck
24	72	10.6	135	1	beta-galactoside-b
25	72	10.6	361	2	probable pectinest
26	71	10.4	135	1	galectin 1 [valida
27	71	10.4	288	2	chromosome partiti
28	71	10.4	392	2	probable myo-inosi
29	71	10.4	396	2	1-deoxy-d-xyulose

30	71	10.4	468	2	AI2811	serine proteinase
31	71	10.4	495	2	B97590	htra protein homol
32	70.5	10.3	688	2	H96681	protein FIE2.10 l
33	70	10.3	303	2	D95011	hypothetical prote
34	70	10.3	320	2	G97882	hypothetical prote
35	69.5	10.2	746	2	T35811	probable phosphodi
36	69	10.1	119	2	G32536	T-cell receptor al
37	69	10.1	165	2	AH2010	molybdopterin conv
38	69	10.1	240	2	AD2534	hypothetical prote
39	69	10.1	285	2	F88281	protein ZK892.1 l
40	69	10.1	297	2	T28073	hypothetical prote
41	69	10.1	310	2	T33972	hypothetical prote
42	69	10.1	544	2	A95591	hypothetical prote
43	68.5	10.0	236	2	AE0167	probable leucyl/ph
44	68.5	10.0	302	2	D82996	probable transcrip
45	68	10.0	323	2	T26796	hypothetical prote
46	68	10.0	409	1	S75020	acetyl-CoA C-acety
47	68	10.0	457	2	T05651	hypothetical prote
48	68	10.0	588	2	A49618	probable ataxia-te
49	67.5	9.9	203	2	D81116	RNA methyltransfer
50	67.5	9.9	261	2	B81904	probable RNA methy
51	67.5	9.9	398	2	AC3578	glycine betaine/l-
52	67.5	9.9	461	1	A46077	steroid hormone re
53	67.5	9.9	554	2	S46346	gag polyprotein -
54	67.5	9.9	790	1	PLPG	plasmin (EC 3.4.21
55	67.5	9.9	1668	1	C69224	cobalamin biosynth
56	67	9.8	244	2	AG1013	probable exported
57	67	9.8	356	2	AG2907	conserved hypothet
58	67	9.8	378	2	G97682	hypothetical prote
59	67	9.8	488	2	G65216	hypothetical 53.4
60	67	9.8	489	2	AG3038	conserved hypothet
61	67	9.8	515	2	D98447	mexT protein (AP11
62	67	9.8	946	2	S27921	nuclear antigen EB
63	67	9.8	1123	1	WMBE7	UL37 protein - hum
64	67	9.8	1339	2	T40245	probable transcrip
65	66.5	9.8	369	2	T30592	conserved hypothet
66	66.5	9.8	686	2	G82130	2,4-dienoyl-CoA re
67	66.5	9.8	746	2	S34656	amine oxidase (cop
68	66	9.7	219	2	AF0639	flagellar basal bo
69	66	9.7	349	2	AF2130	iron(III) dicitrat
70	66	9.7	398	2	D83956	hypothetical prote
71	66	9.7	624	2	H71316	probable penicilli
72	66	9.7	642	2	S58154	hypothetical prote
73	66	9.7	696	2	S61502	flagellar biosynth
74	66	9.7	703	2	B41538	ABC-type transport
75	65.5	9.6	159	2	I84615	coagulation factor
76	65.5	9.6	283	2	T19933	hypothetical prote
77	65.5	9.6	392	2	S56557	hypothetical 41.4K
78	65.5	9.6	395	2	F83873	hypothetical prote
79	65.5	9.6	441	2	AH0603	conserved hypothet
80	65.5	9.6	646	2	AD0912	probable lipoprote
81	65.5	9.6	694	2	A95866	probable transketo
82	65.5	9.6	893	2	AG3022	two component sens
83	65.5	9.6	900	2	C98262	poly-component sens
84	65.5	9.6	2124	2	T28658	polyketide synthas
85	65	9.5	135	1	LNRTGB	beta-galactoside-b
86	65	9.5	300	2	JC4367	NAD(P)-arginine AD
87	65	9.5	303	2	T29513	hypothetical prote
88	65	9.5	384	2	E72507	probable histidin
89	65	9.5	425	2	AI0971	3-deoxy-D-manno-oc
90	65	9.5	437	2	AH2159	hypothetical prote
91	65	9.5	1533	2	A46221	abdominal segment
92	64.5	9.5	240	2	S11293	SA85-1.2 protein -
93	64.5	9.5	284	2	A70976	hypothetical prote
94	64.5	9.5	316	2	AC2500	group 2 sigma 70-t
95	64.5	9.5	392	2	F91290	probable transport
96	64.5	9.5	392	2	H86131	probable transcrip
97	64.5	9.5	466	1	KPHU7	coagulation factor
98	64.5	9.5	585	2	C70341	acetolactate synth
99	64	9.4	206	2	H82100	ribonuclease HII V
100	64	9.4	225	2	AG1000	ribulose-phosphate



Best Local Similarity 27.3%; Pred. No. 0.075;  
Matches 30; Conservative 21; Mismatches 40; Indels 19; Gaps 3;  
  
QY 13 LPOGLSPQGVIIIRGLVLOEPKHTVSLRDOAAHAPVTLRASFAADRTLAWISR----- 65  
DB 46 VPFGI--GVYINRSDINRPLNPAVYTERAIARPLADYGINRVANPTNYLWMLAL 103  
  
QY 66 -----WGQKKLISAPFLFYQPQFFVLLLFQEG---GLKIALNGQGLG 105  
DB 104 LAPTALSMWQLYLAKTGSTLPKRWFGVKVNLNBSGTPFGLATVVVREGIG 153  
  
RESULT 5  
AC1908  
heterocyst-specific glycolipids-directing protein (imported) - Nostoc sp. (strain PCC 7120)  
C;Species: Nostoc sp. PCC 7120  
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C;Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Jul-2004  
C;Accession: AC1908  
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.  
DNA Res. 8, 205-213, 2001  
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena sp. strain PCC 7120  
A;Reference number: AB1807; MUID:21595285; PMID:11759840  
A;Accession: AC1908  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-727 <KUR>  
A;Cross-references: UNIPROT:Q44230; GB:BA000019; PIDN:BAB72770.1; PID:g17130158; GSPDB:G17130158  
A;Experimental source: strain PCC 7120  
A;Genetics: hglK  
  
Query Match 13.9%; Score 94.5; DB 2; Length 727;  
Best Local Similarity 27.3%; Pred. No. 0.075;  
Matches 30; Conservative 21; Mismatches 40; Indels 19; Gaps 3;  
  
QY 13 LPOGLSPQGVIIIRGLVLOEPKHTVSLRDOAAHAPVTLRASFAADRTLAWISR----- 65  
DB 46 VPFGI--GVYINRSDINRPLNPAVYTERAIARPLADYGINRVANPTNYLWMLAL 103  
  
QY 66 -----WGQKKLISAPFLFYQPQFFVLLLFQEG---GLKIALNGQGLG 105  
DB 104 LAPTALSMWQLYLAKTGSTLPKRWFGVKVNLNBSGTPFGLATVVVREGIG 153  
  
RESULT 6  
A55975  
Galectin-8 - rat  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 23-Mar-1995 #sequence\_revision 05-Apr-1995 #text\_change 09-Jul-2004  
C;Accession: A55975  
R;Hadari, Y.R.; Paz, K.; Dekel, R.; Mestrovic, T.; Accili, D.; Zick, Y.  
J. Biol. Chem. 270, 3447-3453, 1995  
A;Title: Galectin-8. A new rat lectin, related to galectin-4.  
A;Reference number: A55975; MUID:95155445; PMID:7852431  
A;Accession: A55975  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-316 <HAD>  
A;Cross-references: UNIPROT:Q62665; GB:U09824; NID:g1717031; PIDN:AAA66359.1; PID:g1717032  
C;Superfamily: lactose-binding lectin L-36  
  
Query Match 12.8%; Score 87.5; DB 2; Length 316;  
Best Local Similarity 23.9%; Pred. No. 0.16;  
Matches 32; Conservative 20; Mismatches 77; Indels 5; Gaps 1;  
  
QY 5 LEVPCSHALPQGLSPGVIIIRGLVLOEPKHTVSLR-----RDQAAHAPVTLRASFAADRT 59  
DB 182 LSLPPEARLNASMGPRIVVKGENVATSFNVDLVAGRSRDIALHNLPLNVKAFVFN 241  
  
QY 60 LAWISRWQKKLISAPFLFYQPQFFVLLLFQEGGLKIALNGQGLGATSMNQAL 119  
DB 184 VIVCNTRKQDNWNGKEERQSA-FPPESGKPFKIQVLVEADHFKVAVDNAHLQYNHRMKNL 242

DB 242 SFLODANGBERNITCFPFSSGMFYFEMIIYCDVREFKVAVGVHSLYKHFRLSSIDT 301  
QY 120 LRISGSVOLYCVHS 133  
DB 302 LAVDGDRLDLVRS 315  
  
RESULT 7  
A28651  
galactose-specific lectin - mouse  
N;Alternate names: carbohydrate-binding protein 35; IGE-binding protein; lectin L-34; Mac-2; Species: Mus musculus (house mouse)  
C;Date: 28-Aug-1989 #sequence\_revision 10-Feb-1995 #text\_change 09-Jul-2004  
C;Accession: S08537; A28651; A37385; A35185  
R;Cherayil, B.J.; Weiner, S.J.; Pillai, S.  
J. Exp. Med. 170, 1959-1972, 1989  
A;Title: The Mac-2 antigen is a galactose-specific lectin that binds IGE.  
A;Reference number: S08537; MUID:90063462; PMID:2584931  
A;Accession: S08537  
A;Status: nucleic acid sequence not shown; translation not shown  
A;Molecule type: mRNA  
A;Residues: 1-264 <CHE>  
A;Cross-references: UNIPROT:P16110; EMBL:X16834; NID:g52986; PIDN:CAA34736.1; PID:g52987  
A;Note: this sequence was submitted to the EMBL Data Library, Oct-1989  
R;Jia, S.; Wang, J.L.  
J. Biol. Chem. 263, 6009-6011, 1988  
A;Title: Carbohydrate binding protein 35. Complementary DNA sequence reveals homology with galectin-8.  
A;Reference number: A28651; MUID:88198129; PMID:3360772  
A;Accession: A28651  
A;Molecule type: mRNA  
A;Residues: 1-3-264 <JIA>  
R;Raz, A.; Pazerini, G.; Carmi, P.  
Cancer Res. 49, 3489-3493, 1989  
A;Title: Identification of the metastasis-associated, galactoside-binding lectin as a chondroitinase B-like protein.  
A;Reference number: A37385; MUID:89275058; PMID:2525069  
A;Accession: A37385  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-3,'T','5-91,'ST','94-109,'SAP',113-264 <RAZ>  
A;Cross-references: GB:X16074; NID:g52850; PIDN:CAA34206.1; PID:g52851  
A;Note: authors translated the codon GAA for residue 219 as Ala, GAC for residue 221 as Cys.  
R;Woo, H.J.; Shaw, L.M.; Messier, J.M.; Mercurio, A.M.  
J. Biol. Chem. 265, 7097-7099, 1990  
A;Title: The major non-integrin laminin binding protein of macrophages is identical to chondroitinase B-like protein.  
A;Reference number: A35185; MUID:90236991; PMID:2332426  
A;Accession: A35185  
A;Molecule type: protein  
A;Residues: 159-163;166-175;214-226 <WOO>  
C;Superfamily: beta-galactoside-binding lectin  
C;Keywords: lectin; phosphoprotein  
  
Query Match 12.5%; Score 85.5; DB 2; Length 264;  
Best Local Similarity 25.4%; Pred. No. 0.21;  
Matches 34; Conservative 22; Mismatches 63; Indels 15; Gaps 4;  
  
QY 5 LEVPCSHALPQGLSPGVIIIRGLVLOEPKHTVSLR---DQAAHAPVTLRASFAADRTLA 61  
DB 128 LTVFYDLPFGVMPRLITMGTVKPNANRIVLDFRRGNDVAFH----FNPFRNENRR 183  
  
QY 62 WI-----SRWGKKLISAPFLFYQPQFFVLLLFQEGGLKIALNGQGLGATSMNQAL 114  
DB 184 VIVCNTRKQDNWNGKEERQSA-FPPESGKPFKIQVLVEADHFKVAVDNAHLQYNHRMKNL 242  
  
QY 115 EQLRELIRSGSVQL 128  
DB 243 REISQLGISGDITL 256  
  
RESULT 8  
A45983  
lactose-binding lectin Mac-2 - mouse  
C;Species: Mus musculus (house mouse)

C;Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 09-Jul-2004  
C;Accession: A45983  
R;Rosenberg, I.M.; Iyer, R.; Cherayil, B.; Chiodino, C.; Pillai, S.  
J. Biol. Chem. 268, 12393-12400, 1993  
A;Title: Structure of the murine Mac-2 gene. Splice variants encode proteins lacking functionally important regions.  
A;Reference number: A45983; MUID:93286070; PMID:8509379  
A;Accession: A45983  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-264 <ROS>  
A;Cross-references: UNIPROT:Q8C253; GB:L08649  
C;Genetics:  
A;Introns: 6/3; 128/3; 158/2; 209/2; 213/3  
C;Superfamily: beta-galactoside-binding lectin

Query Match 12.5%; Score 85.5; DB 2; Length 264;  
Best Local Similarity 25.4%; Pred. No. 0.21;  
Matches 34; Conservative 22; Mismatches 63; Indels 15; Gaps 4;

QY 5 LEVPCSHALPQGLSPQGVIIIVRGVLQEPKHPVTSR---DQAAHAPVTLRASFAADRTLA 61  
Db 128 LTVPYDLPFGGVMRMLITMGTVKPNANRIVLDFRGNDAVFH----FNPRFNNRR 183  
QY 62 WI-----SRWGQKLISAPFLFYQRFVEVLLLFQEGGLKALNGQGLGATSMNQAL 114  
Db 184 VIVCNTKQDNWNGEERQSA-FFESGKPKIQVLVEADHFKVAVNDHALHLLQYNHRMKNL 242  
QY 115 EQLRELRISGSVOL 128  
Db 243 REISQLGISGDITL 256

RESULT 9  
A46631  
lactose-binding lectin L-36 - rat  
N;Alternate names: galectin-4  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
C;Accession: A46631; S69096  
R;Oda, Y.; Herrmann, J.; Gitt, M.A.; Turck, C.W.; Burlingame, A.L.; Barondes, S.H.; Lefebvre, J.A.; et al. J. Biol. Chem. 268, 5929-5939, 1993  
A;Title: Soluble lactose-binding lectin from rat intestine with two different carbohydrate specificities.  
A;Reference number: A46631; MUID:93194902; PMID:8449956  
A;Accession: A46631  
A;Status: preliminary  
A;Molecule type: nucleic acid  
A;Residues: 1-324 <ODA>  
A;Cross-references: UNIPROT:P38552; GB:M73553; NID:g294571; PID:g294572  
A;Experimental source: intestine  
A;Note: sequence extracted from NCBI backbone (NCBIN:128409, NCBIP:128410)  
R;Hardy, F.; Deviller, P.; Louisot, P.; Martin, A.  
FEBS Lett. 359, 169-172, 1995  
A;Title: Purification and characterization of the N-terminal domain of galectin-4 from rat intestine.  
A;Reference number: S69096; MUID:95172227; PMID:7867792  
A;Accession: S69096  
A;Molecule type: protein  
A;Residues: 13-37;44-50,'E',52-66 <TAR>  
C;Superfamily: lactose-binding lectin L-36  
C;Keywords: lectin

Query Match 12.3%; Score 84; DB 2; Length 324;  
Best Local Similarity 26.1%; Pred. No. 0.39;  
Matches 35; Conservative 23; Mismatches 56; Indels 20; Gaps 5;

QY 7 VPCSHALPQGLSPQGVIIIVRGVLQEPKHPVTSR-----DQAAH-----APVTLRASFA 56  
Db 194 VPVVGTGLQGLTARRTIKGYVLPATAKNLIINFKVSGTGDIAFHMNPRIGDCVVRNSYM 253  
QY 57 DRTLAWISRWG--QKLLISAPFLFYQRFVEVLLLFQEGGLKALNGQGLGATSMNQAL 114  
Db 254 N-----GSMGSEERKIPYNP--FGAGQFFDLIRCGTDRFKVPFANGQHLFDPSHRFOAF 305  
QY 115 EQLRELRISGSVOL 128

Db 306 QRVDMLKIKDITL 319  
RESULT 10  
A54740  
interleukin-4-induced transcription factor stat - human  
C;Species: Homo sapiens (man)  
C;Date: 28-Oct-1994 #sequence\_revision 28-Oct-1994 #text\_change 09-Jul-2004  
C;Accession: A54740  
R;Hou, J.; Schindler, U.; Henzel, W.J.; Ho, T.C.; Brasseur, M.; McKnight, S.L.  
Science 265, 1701-1706, 1994  
A;Title: An interleukin-4-induced transcription factor: IL-4 stat.  
A;Reference number: A54740; MUID:94367369; PMID:8085155  
A;Accession: A54740  
A;Status: preliminary; not compared with conceptual translation  
A;Molecule type: mRNA  
A;Residues: 1-848 <HOU>  
A;Cross-references: UNIPROT:P42226  
C;Superfamily: human signal transducer and transcription activator STATSA  
C;Keywords: DNA binding; transcription regulation

Query Match 11.9%; Score 81; DB 2; Length 848;  
Best Local Similarity 29.9%; Pred. No. 2.4;  
Matches 44; Conservative 14; Mismatches 49; Indels 40; Gaps 9;

QY 1 MSPRLEVPCHALP--QGLSPQGVIIIVRGVLQEP-----KHPVTSL- 40  
Db 688 MYPQVYPHSHSIPPYQGLSPESVNVLS-AFQEPHLQMPPLSGQMSLFPDQHPGOLLP 746  
QY 41 ---RQAAHAPVTLRASFAADRTLAWISRWGQKLISAPFLFYQ-----RFEVLLLFQ 91  
Db 747 CQPEHGVSSPDLCS--DVTWV-----EVSCLSQPVTAFPQGTWIGEDIFPPLPPT 798  
QY 92 EGGLE-KLALNGQ-IGATSMNQALEQ 116  
Db 799 EQDLTKLLEGQSGSGSLGAQPLLQ 825

RESULT 11  
galectin-3 - rabbit  
C;Species: Oryctolagus cuniculus (domestic rabbit)  
C;Date: 16-Nov-1995 #sequence\_revision 08-Feb-1996 #text\_change 09-Jul-2004  
C;Accession: JC4300  
R;Gaudin, J.C.; Monsigny, M.; Legrand, A.  
Gene 163, 249-252, 1995  
A;Title: Cloning of the cDNA encoding rabbit galectin-3.  
A;Reference number: JC4300; MUID:96011642; PMID:7590275  
A;Accession: JC4300  
A;Molecule type: mRNA  
A;Residues: 1-242 <GAU>  
A;Cross-references: UNIPROT:P47845; GB:U06470; NID:g606794; PID:AAC48491.1; PID:g606795  
A;Experimental source: vascular smooth muscle cells  
A;Note: The authors translated the codon TTC for residue 155 as Leu  
C;Comment: This protein has the functions on cell adhesion and proliferation. It is a su  
C;Genetics:  
A;Gene: Igals3  
C;Superfamily: beta-galactoside-binding lectin  
C;Keywords: muscle

Query Match 11.8%; Score 80.5; DB 2; Length 242;  
Best Local Similarity 25.0%; Pred. No. 0.65; Indels 19; Gaps 5;

QY 5 LEVPCSHALPQGLSPQGVIIIVRGVLQEPKHPVTSL-----RDQAAHAPVTLRASFAADRT 59  
Db 106 LPVPYDLPFGGVMRMLITMGTV--KFNANRLALDFKRGNDVAFH----FNPRFNN 159  
QY 60 LAWI-----SRWGQKLISAPFLFYQRFVEVLLLFQEGGLKALNGQGLGATSMNQ 112  
Db 160 RRIVCNTKVDNNWNGREER-QTTFPFEIGKPKIQVLVEADHFKVAVNDHALHLLQYNHRM 218

```

Qy      62 WISRWGQKLIASPLFYPQRFVEVLLLFQEGGKLALNGQGLGATSMNQOALQLREL 121
      :||:::||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      252 LNWGAERKSKSNFPAPQIFDLSIRCLGRFKVYANGQHLFDPSHRLSNFGQVDLLE 311
      :||:::||||:||||:||||:||||:||||:||||:||||:||||:
Qy      122 ISGSVOL 128
      :||:::||||:||||:||||:||||:||||:||||:||||:
Db      312 IQGDVTL 318
      :||:::||||:||||:||||:||||:||||:||||:||||:

RESULT 14
A54909
carbohydrate-binding protein CBP30 - hamster
N:Alternate names: S-type animal lectin CBP30
C/Species: Cricetinae Gen. sp. (hamster)
C/Date: 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change 20-Aug-1999
C/Accession: A54909
R:/Mehul, B.; Bawumia, S.; Martin, S.R.; Hughes, R.C.
J. Biol. Chem. 269, 18250-18258, 1994
A:/Title: Structure of baby hamster kidney carbohydrate-binding protein CBP30, an S-type
A:/Reference number: A54909; MUID:94299546; PMID:8027086
A:/Accession: A54909
A:/Status: preliminary
A:/Molecule type: mRNA
A:/Residues: 1-245 <MEH>
A:/Cross-references: GB:X78879; NID:G535082; PIDN:CAA55479.1; PID:G535083
C:/Superfamily: beta-galactoside-binding lectin
C:/Keywords: lectin

Query Match      11.4%; Score 77.5; DB 2; Length 245;
Best Local Similarity 25.0%; Pred. NO. 1.4;
Matches 34; Conservative 23; Mismatches 60; Indels 19; Gaps 5;

Qy      5 LEVPCSHALPQGLSPQGVIVRGLVLOEPKHTVSL-----RDQAAHAPVTLRASFA DRT 59
      |||:::||||:||||:||||:||||:||||:||||:||||:
Db      109 LTPVPKPLAGWMPRLITMGTV--KPNANRIILNFLRGNDIAFH----FNPRFNENN 162
      |||:::||||:||||:||||:||||:||||:||||:||||:
Qy      60 LAWI-----SRWGQKLIASPLFYPQRFVEVLLLFQEGGKLALNGQGLGATSMNQO 112
      :||:::||||:||||:||||:||||:||||:||||:||||:
Db      163 RRIVCNVTKQDNWNGREERQSA-FPFESGRPFKIQLVLEADHFKVAVNDAHLQLYNHRMK 221
      :||:::||||:||||:||||:||||:||||:||||:||||:
Qy      113 ALEQLRELRIISGSVOL 128
      :||:::||||:||||:||||:||||:||||:||||:
Db      222 NLREINQMEISGDI TL 237
      :||:::||||:||||:||||:||||:||||:||||:

RESULT 15

```

N:Alternate names: CBP 35; epithelial-specific lectin 35; galactoside-binding lectin L-2  
C:Species: Homo sapiens (man)  
C>Date: 12-Oct-1990 #sequence\_revision 12-Oct-1990 #text\_change 09-Jul-2004  
C/Accession: A35820; JQ0916; A47473; A36071; A49800  
R:Robertson, M.W.; Albrandt, K.; Keller, D.; Liu, F.T.  
Biochemistry 29, 8093-8100, 1990  
A>Title: Human IgG-binding protein: a soluble lectin exhibiting a highly conserved inter-  
A/Reference number: A35820; MUID:91084480; PMID:2261464  
A/Accession: A35820  
A>Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-250 <ROB>  
A/Cross-references: UNIPROT:PI7931; GB:M57710; NID:g179530; PIDN:AAA35607.1; PID:g179531  
R:Oda, Y.; Leffler, H.; Sakakura, Y.; Kasai, K.; Barondes, S.H.  
Gene 99, 279-283, 1991  
A>Title: Human breast carcinoma cDNA encoding a galactoside-binding lectin homologous to  
A/Reference number: JQ0916; MUID:91216471; PMID:2022338  
A/Accession: JQ0916  
A/Molecule type: mRNA  
A/Residues: 1-250 <ODA>  
A/Cross-references: GB:M36682; NID:g186921; PIDN:AAA36163.1; PID:g186922  
R:Loetz, M.M.; Andrews Jr., C.W.; Korzeilius, C.A.; Lee, E.C.; Steele Jr., G.D.; Clarke, A.  
Proc. Natl. Acad. Sci. U.S.A. 90, 3466-3470, 1993  
A>Title: Decreased expression of Mac-2 (carbohydrate binding protein 35) and loss of its



Db 9 PTLSESLALPE-TKPAQE-YCRGIVTQKPMKXKSTTIQFELAAAINAQVKDKIAYA 66  
Qy 48 PVTURASFADRTLA-----WISWGQKQL-----ISAPFLFYQRFEEVL 87  
Db 67 LPELRCTFGDRSIVFDIADVWKNQLPLDNGGEIADRFDRAPDWLIEIL 114

RESULT 19  
T37216  
beta-galactoside-binding protein GBP - Caenorhabditis elegans  
N:Alternate names: beta-galactoside-binding lectin homolog, 32K  
C:Species: Caenorhabditis elegans  
C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C:Accession: T37216; T26324; A42846; S27798  
R:Arata, Y.; Hirabayashi, J.; Kasai, K.  
J. Biol. Chem. 272, 26669-26677, 1997  
A:Title: Structure of the 32-kDa galectin gene of the nematode Caenorhabditis elegans.  
A:Reference number: Z21635; MUID:97476274; PMID:9334250  
A:Accession: T37216  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-279 <ARA>  
R:Smeye, R.  
submitted to the EMBL Data Library, November 1996  
A:Reference number: Z20198  
A:Accession: T26324  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-279 <WTL>  
A:Cross-references: EMBL:Z82081; PIDN:CAB04959.1; GSPDB:GN00020; CESP:W09H1.6a  
A:Experimental source: clone W09H1  
R:Hirabayashi, J.; Sacoh, M.; Kasai, K.  
J. Biol. Chem. 267, 15485-15490, 1992  
A:Title: Evidence that Caenorhabditis elegans 32-kDa beta-galactoside-binding protein is  
A:Reference number: A42846; MUID:92348399; PMID:1639789  
A:Accession: A42846  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-279 <HIR>  
A:Cross-references: EMBL:M94671; NID:gl56209; PIDN:AAB87718.1; PID:gl56210  
A:Note: sequence extracted from NCBI backbone (NCBIN:109554)  
C:Genetics:  
A:Gene: W09H1.6a  
A:Map position: 2  
A:Introns: 9/3; 201/3  
C:Superfamily: lactose-binding lectin L-36

Query Match 10.6%; Score 72.5; DB 2; Length 279;  
Best Local Similarity 25.6%; Pred. No. 5.4;  
Matches 20; Conservative 15; Mismatches 34; Indels 9; Gaps 2;

Qy 7 VPCSHALPQGLSPGVIIVRGLVLOEPKHPVTSLRDQAA-----HAPVTLRASPADRTLA 61  
Db 11 VPVRSVLQKPEFGQTLIVKGTIDESQRTINLHSTADFSGNDVPLHVSVRFDEGKIV 70

Qy 62 WIS-----RWGQKKLISAP 75  
Db 71 LNSFSNGEWGKEERKSNP 88

RESULT 20  
T26325  
hypothetical protein W09H1.6b - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T26325  
R:Smeye, R.  
submitted to the EMBL Data Library, November 1996  
A:Reference number: Z20198  
A:Accession: T26325  
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA  
A:Residues: 1-285 <WIL>  
A:Cross-references: UNIPROT:O45904; EMBL:Z82081; PIDN:CAB04960.1; GSPDB:GN00020; CESP:W09H1  
A:Experimental source: clone W09H1  
C:Genetics:  
A:Gene: CESP:W09H1.6b  
A:Map position: 2  
A:Introns: 15/3; 207/3  
C:Superfamily: lactose-binding lectin L-36

Query Match 10.6%; Score 72.5; DB 2; Length 285;  
Best Local Similarity 25.6%; Pred. No. 5.5;  
Matches 20; Conservative 15; Mismatches 34; Indels 9; Gaps 2;

Qy 7 VPCSHALPQGLSPGVIIVRGLVLOEPKHPVTSLRDQAA-----HAPVTLRASPADRTLA 61  
Db 17 VPVRSVLQKPEFGQTLIVKGTIDESQRTINLHSTADFSGNDVPLHVSVRFDEGKIV 76

Qy 62 WIS-----RWGQKKLISAP 75  
Db 77 LNSFSNGEWGKEERKSNP 94

RESULT 21  
MNVTIV  
nonstructural protein NS - Toscana virus  
C:Species: Toscana virus  
A:Note: host (mosquito); Homo sapiens (man)  
C>Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 09-Jul-2004  
C:Accession: A38552  
R:Giorgi, C.; Accardi, L.; Nicolletti, L.; Gro, M.C.; Takehara, K.; Hilditch, C.; Morikawa  
Virology 180, 738-753, 1991  
A:Title: Sequences and coding strategies of the S RNAs of Toscana and Rift Valley fever  
A:Reference number: A38552; MUID:91111992; PMID:1846496  
A:Accession: A38552  
A:Molecule type: genomic RNA  
A:Residues: 1-316 <GIO>  
A:Cross-references: UNIPROT:P21699; GB:X53794; NID:g62201; PIDN:CAA37802.1; PID:g62202  
C:Genetics:  
A:Map position: segment S  
C:Superfamily: Toscana virus nonstructural protein NS  
C:Keywords: nonstructural protein

Query Match 10.6%; Score 72.5; DB 1; Length 316;  
Best Local Similarity 27.6%; Pred. No. 6.2;  
Matches 37; Conservative 17; Mismatches 43; Indels 37; Gaps 7;

Qy 10 SHAL-----PQGLSPGVIIVRGLVLOEPKHPVTSLRDQAAHAPVTLRASPADRTLA 64  
Db 63 SHFLIEKEFPVAVLPGFMISAVR-TRLVEP-----TMRELIOE-SIHQJK 104

Qy 65 RWGQKKLISA-----PFLFYQRFEEVL-----FOEGG-----LKLALNGQGLGATSMN 110  
Db 105 RSNKYLLSALRWPTGPTLLEFIDYFEEILFLSEDFDGSIQRYLKLIVKASGLYNSTNE 164

Qy 111 QOALEQURELRISG 124  
Db 165 EQIVEIHRRLVIEG 178

RESULT 22  
H96011  
asparagine synthase (glutamine-hydrolysing) (EC 6.3.5.4) - Sinorhizobium meliloti (strain  
C:Species: Sinorhizobium meliloti  
C>Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 09-Jul-2004  
C:Accession: H96011  
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan  
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001  
A:Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo  
A:Reference number: A95842; MUID:21396508; PMID:11481431  
A:Accession: H96011  
A:Status: preliminary  
A:Molecule type: DNA







A;Title: Human brain lectin (HBL).  
A;Reference number: JC1092  
A;Accession: JC1092  
A;Molecule type: protein  
A;Residues: 2-135 <WAN>  
A;Experimental source: brain  
R;Hirabayashi, J.; Ayaki, H.; Soma, G.I.; Kasei, K.I.  
FEBS Lett. 250, 161-165, 1989  
A;Title: Production and purification of a recombinant human 14 kDa beta-galactoside-binding  
A;Reference number: S04998; MUID:89325559; PMID:2666155  
A;Accession: S04998  
A;Molecule type: protein  
A;Residues: 2,'X','4-16','X','18-40 <HIR>  
A;Experimental source: recombinant protein expressed in E.coli  
R;Castonovo, V.; Luyten, F.; van den Brule, F.; Sobel, M.E.  
Arch. Biochem. Biophys. 297, 132-138, 1992  
A;Title: Identification of a 14-kDa laminin binding protein (HLBP14) in human melanoma c-  
A;Reference number: S24167; MUID:92344405; PMID:1386213  
A;Accession: S24167  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 20-29;50-64;65-74;132-135 <CAS>  
R;Gitt, M.A.; Barondes, S.H.  
Proc. Natl. Acad. Sci. U.S.A. 83, 7603-7607, 1986  
A;Title: Evidence that a human soluble beta-galactoside-binding lectin is encoded by a f  
A;Reference number: A94130; MUID:87016997; PMID:3020551  
A;Accession: C26495  
A;Molecule type: protein  
A;Residues: 22-42,'X','44;66-70,'XX','73,75-78,'Y','80,'X','82;101-102,'X','104-105,'Q','107-1  
R;Hirabayashi J.; Kawasaki, H.; Suzuki, K.; Kasai, K.  
J. Biochem. 101, 987-995, 1987  
A;Title: Further characterization and structural studies on human placenta lectin.  
A;Reference number: A26761; MUID:87279994; PMID:3611046  
A;Accession: A26761  
A;Molecule type: protein  
A;Residues: 70-84,'X','86-87;122-135 <HI2>  
A;Experimental source: Placenta  
C;Comment: The lectin is inactivated upon the formation of disulfide bonds.  
C;Genetics:  
A;Gene: GDB:LGLALS1  
A;Cross-references: GDB:126889; OMIM:150570  
A;Map position: 22q12-22q13.1  
C;Superfamily: beta-galactoside-binding lectin  
C;Keywords: acetylated amino end; glycoprotein; lectin  
F;2-135/Product: beta-galactoside-binding lectin #status experimental <MAT>  
F;2/Modified site: acetylated amino end (Ala) (in mature form) #status predicted  
F;96/Binding site: carbohydrate (Asn) (covalent) #status absent

Query Match 10.4%; Score 71; DB 1; Length 135;  
Best Local Similarity 25.2%; Pred. No. 3.4;  
Matches 31; Conservative 18; Mismatches 64; Indels 10; Gaps 4;

Qy 17 LSPGQVIVRGVLQSPKHPVTSL---RDQAAHAPVTLRASPADRTLAWISR----WGQ 68  
Db 12 LKPGECLEVRGEVAPDAKSFVNLGKDSNNLCLHFNPRNAHGCDANTVCSKDGGAWT 71  
Qy 69 KKLISAPFLFPQRFVEFVLLLFQEGGLKIALNGQGLGATSMNQOALEQLRELIRISGVQL 128  
Db 72 EQR-EAVFPFGQSVAEVCITFDQANLTKVL-PDGVEFKFPNRLNLEAINMAADGDFKI 129  
Qy 129 YCV 131  
Db 130 KCV 132

RESULT 27  
C75570  
Chromosome partitioning protein - Deinococcus radiodurans (strain R1)  
C;Species: Deinococcus radiodurans  
C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C;Accession: C75570  
R;White, O.; Elsen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; M  
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; M

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 266, 1571-1577, 1999  
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A;Reference number: A75250; MUID:20036896; PMID:10567266  
A;Accession: C75570  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-288 <WHI>  
A;Cross-references: UNIPROT:Q9RYD8; GB:AE001865; GB:AE000513; NID:g6457669; PIDN:AAF0960  
A;Experimental source: strain R1  
C;Genetics:  
A;Gene: DR0012  
A;Map position: 1  
C;Superfamily: Bacillus subtilis transport protein spo0J

Query Match 10.4%; Score 71; DB 2; Length 288;  
Best Local Similarity 24.6%; Pred. No. 8;  
Matches 33; Conservative 18; Mismatches 59; Indels 24; Gaps 4;

QY 15 QGLSPGQVIIVRGLVLOEQPKHTVSLRDQ---AAHAPVTLRASFAADRTLAW----- 62  
Db 150 QAVGKGRSTVNNALRLTLPEPVIRALDEGSIASHARAVLTQPEADRLWAFQIRSGRL 209

QY 63 -----ISRWGQKKLISAPFLFYQRPFEVLL--LFOEGGLKLALNGQGLGATSMNQ 111  
Db 210 NVREAALKKEGRDGKQGAPIKVNPFPRAVRQLELDLSRRTGTRVKITGEDKGRVELNY 269

QY 112 QALEQL-RELRISG 124  
Db 270 GSREELDRILQILG 283

RESULT 28  
F72632  
probable myo-inositol-1-phosphate synthase APE1517 - Aeropyrum pernix (strain K1)  
C;Species: Aeropyrum pernix  
C;Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 16-Aug-2004  
C;Accession: F72632  
R;Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; KDNA Res. 6, 83-101, 1999  
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum pernix strain K1  
A;Reference number: A72450; MUID:99310339; PMID:10382966  
A;Accession: F72632  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-392 <RAW>  
A;Cross-references: UNIPROT:Q9YBT2; DBJ:AP000061; NID:g5104821; PIDN:BAAB0516.1; PID:dl  
A;Experimental source: strain K1  
C;Genetics:  
A;Gene: APE1517  
C;Superfamily: Myo-inositol-1-phosphate synthase

Query Match 10.4%; Score 71; DB 2; Length 392;  
Best Local Similarity 27.5%; Pred. No. 11;  
Matches 30; Conservative 14; Mismatches 57; Indels 8; Gaps 2;

QY 15 QGLSPGQVIIVRGLVLOEQPKHTVSLRDQAAHAPVTLRASFAADRTLAWISRWGQK----K 70  
Db 124 QHLRPDVIIVNITTEAEPTLESVRDRVRGLVSQAYALAVLRYAEREARRVALVN 183

QY 71 LISAPFLFYQRPFEVLLLFQEGGLKLALNGQGLGATSMNQALEQLRE 119  
Db 184 AIPAPLANDP-----VLVSMFEAGSLLLGDGATGATPLTADLLEHLAE 228

RESULT 29  
E83188  
1-deoxy-d-xylulose 5-phosphate reductoisomerase PA3650 [imported] - Pseudomonas aeruginosa  
C;Species: Pseudomonas aeruginosa  
C;Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004  
C;Accession: E83188  
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, .; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen.  
A;Reference number: A82950; MUID:20437337; PMID:10984043  
A;Accession: E83188  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-396 <STO>  
A;Cross-references: UNIPROT:Q9KGU6; GB:AE004785; GB:AE004091; NID:g9949809; PIDN:AAG07031  
A;Experimental source: strain PA01  
C;Genetics:  
A;Gene: dxr; PA3650  
C;Superfamily: 1-deoxy-D-xylulose 5-phosphate reductoisomerase

Query Match 10.4%; Score 71; DB 2; Length 396;  
Best Local Similarity 24.1%; Pred. No. 12;  
Matches 32; Conservative 19; Mismatches 50; Indels 32; Gaps 5;

QY 2 SPRELVPCSHA-----LFOGLSPGQVIIVRGLVLOEQ---KHPTVSLRDQAAH----AP 48  
Db 271 NPDMRTPISYAMAWPERIDSGVSLDMFAVGRLDQFQDEQRFPCRLASQAAETGGSAP 330

QY 49 VTLRASFAADRTLAWISRWGQKKLISAPFLFYQRPFEVLLLFQEGGLKLALNGQGLGATS 108  
Db 331 AMLNAA-----NEVAAFLERHIFSDIAVIED-----VLNREAVTAVE 371

QY 109 MNQOALEQLRELR 121  
Db 372 SLDQVLADRRAR 384

RESULT 30  
AI2811  
serine proteinase htrA [imported] - Agrobacterium tumefaciens (strain C58, Dupont)  
C;Species: Agrobacterium tumefaciens  
C;Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 09-Jul-2004  
C;Accession: AI2811  
R;Wood, D.W.; Stetshak, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.; erdge, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClellan; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, E.; ster, B.W.  
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A;Reference number: AB2577; MUID:21608550; PMID:11743193  
A;Accession: AI2811  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-468 <KUR>  
A;Cross-references: UNIPROT:Q8UE46; GB:AE008688; PIDN:AAL42911.1; PID:g17740366; GSPDB:G  
A;Experimental source: strain C58 (Dupont)  
C;Genetics:  
A;Gene: htrA  
A;Map position: circular chromosome  
C;Superfamily: Helicobacter serine proteinase

Query Match 10.4%; Score 71; DB 2; Length 468;  
Best Local Similarity 25.0%; Pred. No. 14;  
Matches 36; Conservative 20; Mismatches 38; Indels 50; Gaps 6;

QY 16 GLSPGQVI-IVRGLVLOEQ-----KHPTVSLRDQAAHAP 48  
Db 309 GIEFGQVTVAVNGLEVHPDALGYRLTTAGIGKSAELTVVEKKEKLTALTAPETAP 368

QY 49 -----VTLRASFAADRTLAWISRWGQKKLISAPFLFYQRPFEVLLL-----FQ 91  
Db 369 RDERLLEGRNPFAGATVANLS-----PKL--ADELRMPSQVTGVVITDVKSGSPAYRVGFQ 422

QY 92 EGGKLKALNGQGLGATSMNQALE 115  
Db 423 PKDVILSLNGADIGSTAAVEKALD 446

Search completed: October 17, 2005, 08:42:51  
Job time : 20 secs

---

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2005, 08:29:48 ; Search time 179 Seconds

(without alignments)  
380.483 Million cell updates/sec

Title: HJACE54

Perfect score: 682

Sequence: 1 msprlevpcshalpqlspg.....leglirelrisgsvqlcyvhs 133

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : UniProt\_03:\*

1: uniprot\_eprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	671	98.4	296	2 Q9NZ03	Q9NZ03 homo sapien
2	671	98.4	336	1 LEGC HUMAN	Q96dt0 homo sapien
3	556	81.5	314	1 LEGC MOUSE	Q91vd1 mus musculus
4	482	70.7	383	1 Q8C3K0	Q8C3K0 mus musculus
5	146.5	21.5	171	2 Q6NV17	Q6NV17 xenopus tro
6	143.5	21.0	171	2 Q6DDR8	Q6DDR8 xenopus lae
7	142.5	20.9	171	2 Q68FJ4	Q68FJ4 xenopus lae
8	134	19.6	315	2 Q7ZTB5	Q7ZTB5 xenopus lae
9	120.5	17.7	172	2 Q6P5T6	Q6P5T6 homo sapien
10	120.5	17.7	172	2 Q8VED9	Q8VED9 mus musculus
11	119	17.4	144	1 LEGS5 RAT	P47967 rattus norv
12	118.5	17.4	172	2 Q9P005	Q9P005 homo sapien
13	109.5	16.1	316	1 LEG8 HUMAN	Q00214 homo sapien
14	108.5	15.9	261	2 Q7ZSY1	Q7ZSY1 xenopus lae
15	102	15.0	322	2 Q9SL83	Q9SL83 mus musculus
16	102	15.0	353	1 LEG9 MOUSE	Q08573 mus musculus
17	99	14.5	323	2 Q6QZP2	Q6QZP2 canis famil
18	98	14.4	322	2 Q6P7Q6	Q6P7Q6 rattus norv
19	98	14.4	326	1 LEG4 MOUSE	Q8K419 mus musculus
20	98	14.4	354	1 LEG9 RAT	P97840 rattus norv
21	97.5	14.3	165	2 Q7TPX9	Q7TPX9 mus musculus
22	97	14.2	308	2 Q7ZTB8	Q7ZTB8 xenopus lae
23	96	14.1	301	1 LEG6 MOUSE	O54891 mus musculus
24	96	14.1	332	2 Q8QGD9	Q8QGD9 gallus gall
25	95.5	14.0	135	1 LEG7 HUMAN	P47929 homo sapien
26	94.5	13.9	727	2 Q44230	Q44230 anabaena sp
27	94.5	13.9	727	2 Q7A2H2	Q7A2H2 anabaena sp
28	94	13.8	204	2 Q7SYPO	Q7SYPO xenopus lae
29	94	13.8	319	2 Q7ZTB6	Q7ZTB6 xenopus lae
30	93.5	13.7	218	2 Q816N9	Q816N9 anopheles s
31	93.5	13.7	316	2 Q6IN24	Q6IN24 rattus norv

32	93	13.6	262	2 Q90713	Q90713 gallus gall
33	91.5	13.4	223	2 Q8C6H0	Q8C6H0 mus musculu
34	91.5	13.4	316	1 LEG8 MOUSE	Q9J115 mus musculu
35	91.5	13.4	359	2 Q9BXC8	Q9BXC8 homo sapien
36	91	13.3	228	2 Q6TGN4	Q6TGN4 brachydanio
37	89.5	13.1	256	2 Q6P8B0	Q6P8B0 xenopus tro
38	88.5	13.0	211	2 Q7QHP5	Q7QHP5 anopheles g
39	88	12.9	334	2 Q6DDQ2	Q6DDQ2 xenopus lae
40	88	12.9	343	2 Q8UW98	Q8UW98 xenopus lae
41	87.5	12.8	300	2 Q96Q87	Q96Q87 homo sapien
42	87.5	12.8	316	1 LEG8 RAT	Q62665 rattus norv
43	87.5	12.8	359	2 Q8TEV1	Q8TEV1 homo sapien
44	87	12.8	356	2 Q6DKI2	Q6DKI2 homo sapien
45	85.5	12.5	263	1 LEG3 MOUSE	P61110 mus musculu
46	85.5	12.5	264	2 Q8C353	Q8C353 mus musculu
47	84	12.3	241	2 Q8NNB5	Q8NNB5 corynebacte
48	84	12.3	324	1 LEG4 RAT	P38552 rattus norv
49	82.5	12.1	135	1 LEG7 MOUSE	O54974 mus musculu
50	82.5	12.1	135	1 LEG7 RAT	P97590 rattus norv
51	82.5	12.1	762	2 Q88MS6	Q88MS6 pseudomonas
52	81.5	12.0	323	1 LEG4 HUMAN	P56470 homo sapien
53	81.5	12.0	340	2 Q8UW99	Q8UW99 xenopus lae
54	81	11.9	847	1 STA6 HUMAN	P42226 homo sapien
55	81	11.9	847	2 Q71UP4	Q71UP4 homo sapien
56	80.5	11.8	136	2 Q9CRB1	Q9CRB1 m mus muscu
57	80.5	11.8	241	1 LEG3 RABIT	P47845 coryctolagus
58	80.5	11.8	558	2 Q7VX36	Q7VX36 bordetella
59	80.5	11.8	558	2 Q7W7Z2	Q7W7Z2 bordetella
60	80.5	11.8	558	2 Q7WLD3	Q7WLD3 bordetella
61	80	11.7	317	2 Q9XSM8	Q9XSM8 sus scrofa
62	80	11.7	349	2 Q9XSM9	Q9XSM9 sus scrofa
63	79	11.6	194	2 Q9BRU6	Q9BRU6 homo sapien
64	78.5	11.5	136	2 Q99ML7	Q99ML7 mus musculu
65	78.5	11.5	197	1 Y105 HAEU	Y105HAEU haemophilus
66	78.5	11.5	261	1 LEG3 RAT	P08699 rattus norv
67	78.5	11.5	323	1 LEG4 PIG	Q29058 sus scrofa
68	78.5	11.5	337	2 Q83WT5	Q83WT5 streptomyce
69	78	11.4	310	2 Q7T118	Q7T118 brachydanio
70	78	11.4	320	2 Q6DHG3	Q6DHG3 brachydanio
71	78	11.4	320	2 Q6TLH6	Q6TLH6 brachydanio
72	78	11.4	320	2 Q7T119	Q7T119 brachydanio
73	77.5	11.4	136	2 Q9D7K4	Q9D7K4 mus musculu
74	77.5	11.4	244	1 LEG3 CRILLO	P47953 cricetus
75	77.5	11.4	249	1 LEG3 HUMAN	P17931 homo sapien
76	77.5	11.4	250	2 Q6FGL0	Q6FGL0 homo sapien
77	77.5	11.4	250	2 Q61BA7	Q61BA7 homo sapien
78	77.5	11.4	328	2 Q9TUB8	Q9TUB8 coryctolagus
79	77.5	11.4	727	2 Q7S218	Q7S218 neurospora
80	77	11.3	320	2 Q8TGN3	Q8TGN3 brachydanio
81	77	11.3	348	2 Q8QGB1	Q8QGB1 oncorhynch
82	77	11.3	530	2 Q8C1W2	Q8C1W2 mus musculu
83	77	11.3	1293	2 Q8MUV2	Q8MUV2 giardia lam
84	77	11.3	1293	2 Q7QSN9	Q7QSN9 giardia lam
85	76.5	11.2	250	2 Q6NVH9	Q6NVH9 homo sapien
86	76.5	11.2	295	1 LEG3 CANPA	P38486 canis famil
87	76	11.1	311	2 Q8WYQ7	Q8WYQ7 homo sapien
88	76	11.1	355	1 LEG9 HUMAN	O00182 homo sapien
89	75	11.0	136	2 Q6J275	Q6J275 ovine aries
90	74.5	10.9	126	2 Q7R579	Q7R579 giardia lam
91	74.5	10.9	277	2 Q9NJV1	Q9NJV1 haemonchus
92	74.5	10.9	283	2 Q86FX2	Q86FX2 haemonchus
93	74.5	10.9	283	2 Q9NUV0	Q9NUV0 haemonchus
94	74.5	10.9	373	2 Q9DCS3	Q9DCS3 mus musculu
95	74.5	10.9	373	2 Q99L39	Q99L39 mus musculu
96	74.5	10.9	555	2 Q64WR0	Q64WR0 bacteroides
97	74.5	10.9	703	2 Q8DASS	Q8DASS vibrio vuln
98	74	10.9	209	2 Q9KHV9	Q9KHV9 rhizobium m
99	74	10.9	652	2 Q6N349	Q6N349 rhodospheo
100	73.5	10.8	283	1 LEG1_HAECO	O44126 haemonchus

ALIGNMENTS

```
RESULT 1
Q9NZ03 ID Q9NZ03 PRELIMINARY; PRT; 296 AA.
AC Q9NZ03;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Galectin-related inhibitor of proliferation isoform b.
GN Name=GRIP1;
OS Homo sapiens (Human).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Metazoa; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A. (ISOFORM A).
RC TISSUE=Liver;
RA Yang R.-Y., Hsu D.K., Ni J., Yu L., Liu F.-T.;
RA Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the galectin (galactin/S-lectin) family.
DR EMBL; AF222694; AAF34676.1; -.
DR HSP; P17931; IAK3.
DR GO; GO:0005529; F:sugar binding; IEA.
DR InterPro; IPR008985; ConA like lec_gl.
DR InterPro; IPR001079; Galectin.
DR InterPro; IPR00504; RNA_rec_mot.
DR Pfam; PF00337; Gal-bind_lectin; 1.
DR SMART; SM00276; GLECT; 1.
DR PROSITE; PS00030; RRM_RNP_1; UNKNOWN_1.
KW Galectin; Lectin.
SQ SEQUENCE 296 AA; 33257 MW; 1A160DA2E9A48ACC CRC64;

Query Match 98.4%; Score 671; DB 2; Length 296;
Best Local Similarity 99.2%; Pred. No. 2.4e-64;
Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSRLVPCSHALPQGLSPQVIVRGLVLPQEPKFTVSRDQAAHAPVTLRASFADRTL 60
Db 164 MSRLVPCSHALPQGLSPQVIVRGLVLPQEPKFTVSRDQAAHAPVTLRASFADRTL 223
Qy 61 AWISRWGKKLISAPFLFYQRFVFFVLLFQEGGLKALNGQGLGATSMNQALEQREL 120
Db 224 AWISRWGKKLISAPFLFYQRFVFFVLLFQEGGLKALNGQGLGATSMNQALEQREL 283
Qy 121 RIGSVQLYCVHS 133
Db 284 RIGSVQLYCVHS 296

RESULT 2
LEGC_HUMAN STANDARD; PRT; 336 AA.
AC Q96DT0; Q96DS9; Q96PR9; Q9H258; Q9H259; Q9NZ02;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Galectin-12 (Galectin-related inhibitor of proliferation).
GN Name=LGALS12; Synonyms=GRIP1;
OS Homo sapiens (Human).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Metazoa; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A. (ISOFORMS A, B, C AND D), AND CHARACTERIZATION.
RC TISSUE=Adipose tissue;
RX MEDLINE=21432969; PubMed=11435439; DOI=10.1074/jbc.M105097200;
RA Hotta K., Funahashi T., Matsukawa Y., Takahashi M., Nishizawa H.,
RA Kishida K., Matsuda M., Kuriyama H., Kihara S., Nakamura T.,
RA Tochino Y., Bodkin N.L., Hansen B.C., Matsuzawa Y.;
RT "Galectin-12, an adipose-expressed galectin-like molecule possessing
RT apoptosis-inducing activity."
RL J. Biol. Chem. 276:34089-34097(2001).
[2]

SEQUENCE FROM N.A. (ISOFORMS B; E AND F).
TISSUE=Retina;
MEDLINE=21283005; PubMed=11283015; DOI=10.1074/jbc.M010914200;
Yang R.-Y., Hsu D.K., Yu L., Ni J., Liu F.-T.;
"Cell cycle regulation by galectin-12, a new member of the galectin
superfamily."
J. Biol. Chem. 276:20252-20260(2001).
[3]
RP SEQUENCE FROM N.A. (ISOFORM A).
RC TISSUE=Blood;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
-1- FUNCTION: Binds lactose. May participate in the apoptosis of
adipocytes.
-1- SUBCELLULAR LOCATION: Nuclear.
-1- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=6;
Name=A;
IsoId=Q96DT0-1; Sequence=Displayed;
Name=B; Synonyms=GRIP1a;
IsoId=Q96DT0-2; Sequence=VSP_003100;
Name=C;
IsoId=Q96DT0-3; Sequence=VSP_003102;
Name=D;
IsoId=Q96DT0-4; Sequence=VSP_003100, VSP_003102;
Name=E; Synonyms=1;
IsoId=Q96DT0-5; Sequence=VSP_003099;
Name=F; Synonyms=2;
IsoId=Q96DT0-6; Sequence=VSP_003099, VSP_003101;
-1- TISSUE SPECIFICITY: Not widely expressed. Predominantly expressed
in adipose tissue.
-1- DOMAIN: Contains two homologous but distinct carbohydrate-binding
domains.
-1- SIMILARITY: Belongs to the galectin (galectin/S-lectin) family.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
EMBL; AF244974; AAK77328.1; -.
EMBL; AF244975; AAK77329.1; -.
EMBL; AF244976; AAK77330.1; -.
EMBL; AF244977; AAK77331.1; -.
EMBL; AF310686; AAG40863.1; -.
EMBL; AF310687; AAG40864.1; -.
EMBL; AF222695; AAF34677.1; -.
EMBL; BC028222; AAH28222.1; -.
HSSP; P17931; IAK3.
DR Genew; HGNC:15788; LGALS12.
DR H-InvDB; HIX0009737; -.

```

```

DR MIM; 606096; --
DR InterPro; IPR008985; ConA-like lec_gl.
DR InterPro; IPR001079; Galectin.
DR Pfam; PF00337; Gal-bind lectin; 1.
DR SMART; SM00276; GLBCT; 1
DR PROSITE; PS00309; GALACTIN; FALSE_NEG.
KW Alternative splicing; Apoptosis; Galectin; Lectin; Nuclear protein;
KW Repeat.
FT DOMAIN 47 183 Galactin 1.
FT DOMAIN 194 336 Galactin 2.
FT VARSPLIC 1 22 Missing (in isoform B and isoform F).
FT VARSPLIC 1 61 Missing (in isoform B and isoform D).
FT VARSPLIC 74 74 Missing (in isoform F).
FT VARSPLIC 200 208 Missing (in isoform C and isoform D).
FT VARSPLIC 336 AA; 37541 MW; F35D0A0CE503E795 CRC64;
SQ SEQUENCE 98.4%; Score 671; DB 1; Length 336;
Query Match 99.2%; Pred. No. 2.8e-64;
Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy 1 MSPRLVPCSHALPQGLSPGVIIIRGLVLPQEPKHTVSLRDQAHAHPVTLRASFADRTL 60
Db 204 MSPRLVPCSHALPQGLSPGVIIIRGLVLPQEPKHTVSLRDQAHAHPVTLRASFADRTL 263
Oy 61 AWISRWGOKLISAPFLYFYPORFVLLFOEGGLKALNGOGLGATSMNQALEQREL 120
Db 264 AWISRWGOKLISAPFLYFYPORFVLLFOEGGLKALNGOGLGATSMNQALEQREL 323
Oy 121 RISGSVOLYCVHS 133
Db 324 RISGSVOLYCVHS 336

RESULT 3
LEGC MOUSE STANDARD; PRT; 314 AA.
ID Q91VD1; Q8CC44; Q8K2L7; Q9UKX2;
AC 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Galectin-12.
GN Name=lgals12;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND CHARACTERIZATION.
RC STRAIN=C57BL/KsK; TISSUE=adipose tissue;
RX MEDLINE=21423969; PubMed=11435439; DOI=10.1074/jbc.M105097200;
RA Hotta K., Funahashi T., Matsukawa Y., Takahashi M., Nishizawa H.,
RA Kishida K., Matsuda M., Kuriyama H., Kihara S., Nakamura T.,
RA Tochino Y., Bodkin N.L., Hansen B.C., Matsuzawa Y.;
RT "Galectin-12, an adipose-expressed galectin-like molecule possessing
RT apoptosis-inducing activity.";
RL J. Biol. Chem. 276:34089-34097 (2001).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC STRAIN=C57BL/6J; TISSUE=Spleen;
RX MEDLINE=21283005; PubMed=11283015; DOI=10.1074/jbc.M010914200;
RA Yang R.-Y., Hsu D.K., Yu L., Ni J., Liu F.-T.;
RT "Cell cycle regulation by galectin-12, a new member of the galectin
RT superfamily.";
RL J. Biol. Chem. 276:20252-20260 (2001).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC STRAIN=C57BL/6J; TISSUE=Colon;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,

```

```

RA Nikaïdo I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yegi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojohori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimonard S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempie C.A., Setou M., Shimada K.,
RA Sultana R., Takemura Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Varadarajan L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavalan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RN Nature 420:563-573 (2002).
RP [4]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.F., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haeh F.,
RA Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McSwain P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Pahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RP [5]
RP FUNCTION: Binds lactose. May participate in the apoptosis of
RP adipocytes.
RP [6]
RP SUBCELLULAR LOCATION: Nuclear.
RP [7]
RP ALTERNATIVE PRODUCTS:
RP Event=Alternative splicing; Named isoforms=2;
RP Name=1;
RP IsoId=Q91VD1-1; Sequence=Displayed;
RP Name=2;
RP IsoId=Q91VD1-2; Sequence=VSP_010321;
RP [8]
RP DOMAIN: Contains two homologous but distinct carbohydrate-binding
RP domains.
RP [9]
RP SIMILARITY: Belongs to the galectin (galactin/s-lectin) family.
RP [10]
RP This SWISS-PROT entry is copyright. It is produced through a collaboration
RP between the Swiss Institute of Bioinformatics and the EMBL outstation -
RP the European Bioinformatics Institute. There are no restrictions on its
RP use by non-profit institutions as long as its content is in no way
RP modified and this statement is not removed. Usage by and for commercial
RP entities requires a license agreement (See http://www.isb-sib.ch/announce/
RP or send an email to license@isb-sib.ch)
RP [11]
RP EMBL; AF244979; AAK77327.1; -

```

```

DR EMBL; AF244978; AAK77326.1; -.
DR EMBL; AF223223; AAF34682.1; -.
DR EMBL; AK033355; BAC28345.1; -.
DR EMBL; BC030890; AAH30890.1; -.
DR HSSP; P17931; 1A3K.
DR MGD; MGI:1929094; Lgale12.
DR InterPro; IPR008985; ConA_like_lect_gl.
DR InterPro; IPR001079; Galectin.
DR Pfam; PF00337; Gal-bind_lectin; 1.
DR SMART; SM00276; GLECT; 1.
KW Alternative splicing; Apoptosis; Galectin; Lectin; Nuclear protein;
FT Repeat.
FT DOMAIN 25 161 Galaptin 1.
FT DOMAIN 172 314 Galaptin 2.
FT VARSPDIC 167 179 Missing (in isoform 2).
FT /Frid=VSP_010321.
FT CONFLICT 228 228 V -> A (in Ref. 1).
FT CONFLICT 273 273 E -> K (in Ref. 4).
SQ SEQUENCE 314 AA; 35460 MW; 9C9D33A7583DD09D CRC64;

Query Match 81.5%; Score 556; DB 1; Length 314;
Best Local Similarity 81.7%; Pred. No. 8e-52;
Matches 107; Conservative 12; Mismatches 12; Indels 0; Gaps 0;

QY 2 SPRLVPCSHALPQGLSPGQVIVRGLVLOEPKHPTVSLRDQAHAHAPVTLRASFAADRTLA 61
DB 183 SPRLVPCSHALPRLGLWPGQVIVRGLVLEKPEKDFTLSDKDTTHVPTVTLRASFTDRTLA 242

QY 62 WISRWGOKLISAPFLFYPPQRFVEVLLLFQEGGLKALNQGGLGATSMNQALEQLELR 121
DB 243 WVSWSGRKKLISAPFLFHPQRFVEVLLLCQEGGLKALNQGGLGATSLDQKALEQLELR 302

QY 122 ISGSVQLYCVH 132
DB 303 ISGNVHLVYCVH 313

RESULT 4
ID Q8C3K0 PRELIMINARY; PRT; 383 AA.
AC Q8C3K0;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 25, Last annotation update)
DE Mus musculus 10 days lactation, adult female mammary gland cDNA, RIKEN
DE full-length enriched library, clone:D73000J21 product:lectin,
DE galactose binding, soluble 12, full insert sequence.
GN Name=Lgale12;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Mammary gland;
RX MEDLINE=92279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Mammary gland;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Mammary gland;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
```

```

RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Mammary gland;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Mammary gland;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamanoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Mammary gland;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases
CC -!- SIMILARITY: Belongs to the galectin (galactin/S-lectin) family.
DR EMBL; AK085660; BAC39498.1; -.
DR HSSP; P17931; 1A3K.
DR MGD; MGI:1929094; Lgals12.
DR GO; GO:0005739; C:mitochondrion; IDA.
DR GO; GO:0005634; C:nucleus; IDA.
DR InterPro; IPR008985; ConA_like_lect_gl.
DR InterPro; IPR001079; Galectin.
DR Pfam; PF00337; Gal-bind_lectin; 1.
DR SMART; SM00276; GLECT; 1.
KW Galectin; Lectin.
SQ SEQUENCE 383 AA; 43036 MW; B7C10FA5746977DF CRC64;

Query Match 70.7%; Score 482; DB 2; Length 383;
Best Local Similarity 80.9%; Pred. No. 1.1e-43;
Matches 93; Conservative 11; Mismatches 11; Indels 0; Gaps 0;

QY 2 SPRLVPCSHALPQGLSPGQVIVRGLVLOEPKHPTVSLRDQAHAHAPVTLRASFAADRTLA 61
DB 183 SPRLVPCSHALPRLGLWPGQVIVRGLVLEKPEKDFTLSDKDTTHVPTVTLRASFTDRTLA 242

QY 62 WISRWGOKLISAPFLFYPPQRFVEVLLLFQEGGLKALNQGGLGATSMNQALEQ 116
DB 243 WVSWSGRKKLISAPFLFHPQRFVEVLLLCQEGGLKALNQGGLGATSLDQKALEQ 297

RESULT 5
ID Q6NV17 PRELIMINARY; PRT; 171 AA.
AC Q6NV17;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein MGC76272.
```



GN Name=MGC76272;  
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=83364;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RX MEDLINE=22389257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smalish D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences".  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RA Klein S., Gerhard D.S.;  
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: Belongs to the galectin (galactin/S-lectin) family.  
DR GO:0005529; F:sugar binding; IEA.  
DR InterPro: IPR008985; ConA like lec\_gl.  
DR InterPro: IPR001079; Galectin.  
DR Pfam: PF00337; Gal-bind lectin; 1.  
DR SMART: SM00276; GLECT; 1.  
DR Galectin; Hypothetical protein; Lectin.  
SQ SEQUENCE 171 AA; 19111 MW; 39CF31F7C65DA040 CRC64;  
Query Match 21.5%; Score 146.5; DB 2; Length 171;  
Best Local Similarity 31.6%; Pred. No. 1.2e-07;  
Matches 43; Conservative 27; Mismatches 55; Indels 11; Gaps 5;  
QY 3 PRLEVP-CSHALPQGLSPGVIIIVGLVLPQEPKPTVSL-----RDQAAHAPVTLRASFA 56  
DB 33 PRLTVPCGH-IGKGMFGKILINGVNLPEKSPDIRLTCGSDPAADIAELRAEFA 91  
QY 57 DRTL---AWIS-RWGQKLIAPFLFYQRFQFVLLFQEGGLKALNGOGLGATSMNQ 112  
DB 92 DKQLLRNACVSGKMGEESSAIPYFPFIADQPRVRIHCEHPRLRIFVDGHLDFYHRVE 151  
QY 113 ALEQLRELRLISGSVQL 128  
DB 152 TSLAINTIKINGDLQL 167  
RESULT 6  
Q6DDR8 PRELIMINARY; PRT; 171 AA.  
AC Q6DDR8;  
DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE MGC82421 protein.  
GN Name=MGC82421;  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;  
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
RA Richardson P.;  
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
Initiative".  
RL Dev. Dyn. 225:384-391 (2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RX PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smalish D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences".  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RA Klein S., Gerhard D.S.;  
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: Belongs to the galectin (galactin/S-lectin) family.  
DR EMBL: BC077459; AAH77459.1; -.  
DR GO:0005529; F:sugar binding; IEA.  
DR InterPro: IPR008985; ConA like lec\_gl.  
DR InterPro: IPR001079; Galectin.  
DR Pfam: PF00337; Gal-bind lectin; 1.  
DR SMART: SM00276; GLECT; 1.  
DR Galectin; Lectin.  
SQ SEQUENCE 171 AA; 19233 MW; 7FE19E3C7471D402 CRC64;  
Query Match 21.0%; Score 143.5; DB 2; Length 171;  
Best Local Similarity 30.9%; Pred. No. 2.5e-07;  
Matches 42; Conservative 29; Mismatches 54; Indels 11; Gaps 5;  
QY 3 PRLEVP-CSHALPQGLSPGVIIIVGLVLPQEPKPTVSL-----RDQAAHAPVTLRASFA 56  
DB 33 PRLTVPCGH-IGKGMFGKILINGVNLPEKSPDIRLTCGSDPAADIAELRAEFS 91  
QY 57 DRTL---AWIS-RWGQKLIAPFLFYQRFQFVLLFQEGGLKALNGOGLGATSMNQ 112  
DB 92 DKQFLRNARVSGKMGEESSAIPYFPFIADQPRVRIHCEHPRLRIFVDGHLDFYHRVE 151  
QY 113 ALEQLRELRLISGSVQL 128  
DB 152 TSLAINTIKINGDLQL 167  
RESULT 7  
Q68FJ4 PRELIMINARY; PRT; 171 AA.  
ID Q68FJ4  
AC Q68FJ4;  
DT 25-OCT-2004 (TrEMBLrel. 28, Created)

```

DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RC TISSUE=Eve;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RX PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.J., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RA Klein S., Gerhard D.S.;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the galectin (galaptin/s-lectin) family.
DR EMBL: BC079769; AAH79769.1; -.
DR InterPro; IPR008985; ConA_like_lect_gl.
DR InterPro; IPR001079; Galectin.
DR Pfam; PF00337; Gal-bind_lectin; 1.
DR SMART; SM00276; GLEGT; 1.
KW Galectin; Hypothetical protein; Lectin.
SQ SEQUENCE 171 AA; 19184 MW; 6F57F840EE981E55 CRC64;

Query Match 20.9%; Score 142.5; DB 2; Length 171;
Best Local Similarity 30.9%; Pred. No. 3.3e-07;
Matches 42; Conservative 27; Mismatches 56; Indels 11; Gaps 5;

Qy 3 PRLEVP-CSHALPGQLSPGQVIIVRGVLQEPKHTVSL-----RDOAAHAPVTLRASFA 56
Db 33 PRUTVPFCGH-IKGLRPGKILIMGVINLEPKSFRLTCGDSEDPAADIAELRAEFA 91

Qy 57 DRTL---AWIS-RWGOKKLISAPFLFYQRFVFFVLLFQEGGLKALNGOGLGATSMNQ 112
Db 92 DKQFLRNACVSGKNGEESAIPYFPFTADQPRVEIHCEHPQRIIVDGHQLFDFYHRE 151

Qy 113 ALQOLRELISGVSQVL 128
Db 152 SLSAINTIKNGDIQL 167

RESULT 8

```

```

Q7ZTB5
ID Q7ZTB5 PRELIMINARY; PRT; 315 AA.
AC Q7ZTB5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Galectin family xgalectin-VIIIA.
DE Name=xgalectin-VIIIA;
GN Name=xgalectin-VIIIA;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22552404; PubMed=12538594; DOI=10.1074/jbc.M209008200;
RA Shoji H., Nishi N., Hirashima M., Nakamura T.;
RT "Characterization of the Xenopus galectin family. Three structurally
RT different types as in mammals and regulated expression during
RT embryogenesis.";
RL J. Biol. Chem. 278:12285-12293(2003).
CC -!- SIMILARITY: Belongs to the galectin (galaptin/s-lectin) family.
DR EMBL: AB080021; BAC55887.1; -.
DR HSSP; P17931; 1A3K.
DR GO; GO:0005529; F.sugar binding; IEA.
DR InterPro; IPR008985; ConA_like_lect_gl.
DR InterPro; IPR001079; Galectin.
DR Pfam; PF00337; Gal-bind_lectin; 2.
DR SMART; SM00276; GLEGT; 2.
KW Galectin; Lectin.
SQ SEQUENCE 315 AA; 35895 MW; 3A42275840579A9A CRC64;

Query Match 19.6%; Score 134; DB 2; Length 315;
Best Local Similarity 25.9%; Pred. No. 5.6e-06;
Matches 35; Conservative 30; Mismatches 64; Indels 6; Gaps 2;

Qy 3 PR-LEVPCSHALPGQLSPGQVIIVRGVLQEPKHTVSL-----RDOAAHAPVTLRASFA 56
Db 178 PRNFTPTTCTCLSPILPGKTMVKGVLNKRFAIDLPHGSKDIALHLNPMKERVF 237

Qy 57 DRTLAWISRWGOKKLISAPFLFYQRFVFFVLLFQEGGLKALNGOGLGATSMNQALEQ 116
Db 238 VRNTYLRSGWGEKEKQLDFFPCFPMVFETLIICDLQQRVAVNGVHLLYKHFRLDNLK 297

Qy 117 LRELRISGVSQVLXCV 131
Db 298 INEVSVNGDIQLHDV 312

RESULT 9
Q6P5T6
ID Q6P5T6 PRELIMINARY; PRT; 172 AA.
AC Q6P5T6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pooled;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.J., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

```



[illegible]

OY 113 ALEOLRELRISGSVOL 128  
 DB 153 TLSAIDRIKINGDLQI 168

RESULT 13  
 LEG8 HUMAN  
 ID \_LEG8 HUMAN STANDARD; PRT; 316 AA.  
 AC 000214; 015215; Q96B92; Q9H584; Q9H585; Q9UEZ6; Q9UP32; Q9UP33;  
 AC Q9UP34;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 25-JAN-2005 (Rel. 46, Last annotation update)  
 DE Galactin-8 (Gal-8) (prostate carcinoma tumor antigen 1) (PCTA-1) (Po66  
 DE carbohydrate-binding protein) (Po66-CBP).  
 GN Name=LGALS8;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Prostate;  
 RC MEDLINE=96293510; PubMed=8692978; DOI=10.1073/pnas.93.14.7252;  
 RA Su Z.-Z., Lin J., Shen R., Fisher P.B., Goldstein N.I., Fisher P.B.;  
 RT "Surface-epitope mapping and expression cloning identifies the human  
 RT prostate carcinoma tumor antigen gene PCTA-1 a member of the galectin  
 RT gene family.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 93:7252-7257(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Hippocampus;  
 RA Hadari Y.R., Eissenstein M., Zakut R., Zick Y.;  
 RT "Galectin-8: on the road from structure to function.";  
 RL Trends Glycosci. Glycotechnol. 9:103-112(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.; AND ALTERNATIVE SPLICING.  
 RC TISSUE=Lung carcinoma;  
 RA Brichory F., Biron N., Desrues B., Bourguet P., Le Pennec J.P.,  
 RA Dazord L.;  
 RT "Molecular cloning of a beta-galactoside-binding lectin related to  
 RT galectin-8 and identified in human lung carcinoma.";  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Maier C., Haeussler J., Roessch K., Moschgath E., Vogel W.;  
 RT "Genomic organization and expression of the human galectin-8 gene.";  
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20438187; PubMed=10980616; DOI=10.1038/sj.onc.1203767;  
 RA Gopalakrishnan R.V., Roberts T., Tuli S., Kang D., Christiansen K.A.,  
 RA Fisher P.B.;  
 RT "Molecular characterization of prostate carcinoma tumor antigen-1,  
 RT pCTA-1, a human galectin-8 related gene.";  
 RL Oncogene 19:4405-4416(2000).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Colorectal carcinoma;  
 RA Lahm H., Siebert H.-C., Andre S., Hoefflich A., Diehl D., Sordat B.,  
 RA Kaltner H., Wolf E., Gabius H.-J.;  
 RT "Coca (Colorectal carcinoma-derived) galectin-8 variant I full-length  
 RT cDNA from a human colorectal carcinoma cell line.";  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RA Cobley V.;  
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
 RN [8]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
 RC TISSUE=Brain, and Skin;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L.H., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903(2002).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=3;  
 CC Name=1; Synonyms=1;  
 CC IsoId=000214-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=000214-2; Sequence=VSP\_003094;  
 CC Name=3;  
 CC IsoId=000214-3; Sequence=VSP\_003095;  
 CC -!- TISSUE SPECIFICITY: Ubiquitous. Selective expression by prostate  
 CC carcinomas versus normal prostate and benign prostatic  
 CC hypertrophy.  
 CC -!- DOMAIN: Contains two homologous but distinct carbohydrate-binding  
 CC domains.  
 CC -!- SIMILARITY: Belongs to the galectin (galactin/s-lectin) family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; L78132; AAB51605.1; -;  
 DR EMBL; X91790; CAA62904.1; ALT\_INIT.  
 DR EMBL; AF074000; AAD45402.1; -;  
 DR EMBL; AF074001; AAD45403.1; -;  
 DR EMBL; AF074002; AAD45404.1; -;  
 DR EMBL; AF193806; AAF19370.1; ALT\_INIT.  
 DR EMBL; AF193806; AAF19370.1; JOINED.  
 DR EMBL; AF342815; AAK16735.1; ALT\_INIT.  
 DR EMBL; AL136105; CAC15946.1; ALT\_INIT.  
 DR EMBL; AL136105; CAC15947.1; ALT\_INIT.  
 DR EMBL; BC015818; AAHL15818.1; -;  
 DR EMBL; BC016486; AAHL16486.2; -;  
 DR HSSP; P17931; 1A3K.  
 DR Genew; HGNC:6569; LGALS8.  
 DR H-InvDB; HIX0001724; -;  
 DR MiW; 606099; -;  
 DR GO; GO:0005615; C:extracellular space; TAS.  
 DR GO; GO:0005529; F:sugar binding; TAS.  
 DR InterPro; IPR008985; ConA like lec\_gl.  
 DR InterPro; IPR01079; Galectin.  
 DR Pfam; PF00337; Gal-bind lectin; 2.  
 DR PROSITE; PS00309; GALACTIN; 1.  
 DR Alternative splicing; Antigen; Galectin; Lectin; Polymorphism; Repeat.  
 FT DOMAIN 1 153  
 FT Linker.  
 FT DOMAIN 154 184  
 FT DOMAIN 185 316  
 FT SITE 248 254  
 FT VARSPLIC 182 182  
 FT L -> LPSNRGDSKIAPRTVYTKSDSTVNHILCTCTKIP  
 FT PMNYVSK (in isoform 2).





```

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanches A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalish D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Mammary tumor;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the galectin (galactin/S-lectin) family.
CC EMBL; BC003754; AA03754.1; -.
DR HSSP; P17931; 1A3K.
DR MGI; 109496; Lgals9.
DR GO; GO:0005529; F:sugar binding; IEA.
DR InterPro; IPR008985; ConA_like_lect_gl.
DR InterPro; IPR001079; Galectin.
DR SMART; SM00276; GLECT; 2.
DR PROSITE; PS00309; GALAPTIN; 2.
KW Galectin; Lectin.
SQ SEQUENCE 322 AA; 36545 MW; 032D77400737562E CRC64;

Query Match 15.0%; Score 102; DB 2; Length 322;
Best Local Similarity 24.4%; Pred. No. 0.017;
Matches 32; Conservative 27; Mismatches 68; Indels 4; Gaps 2;

Qy 2 SPRLEVPCHALPQGLSPGVIIVRGLVLPQPKHTVSLR---DOAAHAPVTLRASFAADR 58
Db 187 TPATYPTPTIPNGLYPSKSMISGNVLPDTRPHNLRCGGDIAPHNLPRFNENAVVR 246

Qy 59 TLAWISRWGQ-KKLISAPFLFPQRFVLLFQEGGLKALNGGLGATSMNQALRQL 117
Db 247 NTQINNSWGQERSLLGRMPFSRGSFVSFWIICGHCXKVAVNGQHMCEYVHRLKNLQDI 306

Qy 118 RELRISGSVOL 128
Db 307 NTLVAGDIQL 317

RESULT 16
LEG9 MOUSE
ID _LEG9 MOUSE STANDARD; PRT; 353 AA.
AC O08573; O08572;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE Galectin-9
GN Name=Lgals9;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CD-1; TISSUE=Kidney, and Small intestine;
RX MEDLINE=97190351; PubMed=9038233; DOI=10.1074/jbc.272.9.6078;
RA Wada J., Kanwar Y.S.;
RT "Identification and characterization of galectin-9, a novel beta-
RL galactoside-binding mammalian lectin.";
RL J. Biol. Chem. 272:6078-6086(1997).
[2]
RN CHARACTERIZATION.

```

```

RX MEDLINE=97298141; PubMed=9153289;
RA Wada J., Ota K., Kumar A., Wallner E.I., Kanwar Y.S.;
RT "Developmental regulation, expression, and apoptotic potential of
RT galectin-9, a beta-galactoside binding lectin.";
RL J. Clin. Invest. 99:2452-2461(1997).
CC -!- FUNCTION: Binds galactosides. May play a role in thymocyte-
CC epithelial interactions relevant to the biology of the thymus.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic. May also be secreted by a non-
CC classical secretory pathway.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Comment=Additional isoforms seem to exist;
CC Name=Long; IsoId=O08573-1; Sequence=Displayed;
CC Name=Short;
CC IsoId=O08573-2; Sequence=VSP_003097;
CC -!- TISSUE SPECIFICITY: Accentuated expression in liver and thymus of
CC embryo, detected in embryonic heart, brain, lung, liver, and
CC kidney. Highly expressed in adult thymus, small intestine, and
CC liver, and to a lesser extent in lung, kidney, spleen, cardiac,
CC and skeletal muscle. Barely detectable in brain and reticulocyte.
CC The long form is expressed exclusively in the small intestine.
CC -!- DEVELOPMENTAL STAGE: The expression increased with successive
CC stages of embryonic development.
CC -!- DOMAIN: Contains two homologous but distinct carbohydrate-binding
CC domains.
CC -!- SIMILARITY: Belongs to the galectin (galactin/S-lectin) family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.iesb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; U55061; AB51190.1; -.
CC EMBL; U55060; AB51189.1; -.
CC HSSP; P17931; 1A3K.
CC MGI; 109496; Lgals9.
CC InterPro; IPR008985; ConA_like_lect_gl.
CC InterPro; IPR001079; Galectin.
CC Pfam; PF00337; Gal-bind lectin; 2.
CC PROSITE; PS00309; GALAPTIN; 2.
KW Alternative splicing; Galectin; Lectin; Repeat.
FT DOMAIN 1 147 Galaptin 1.
FT DOMAIN 148 204 Linker.
FT DOMAIN 205 353 Galaptin 2.
FT SITE 81 87 Beta-galactoside binding 1 (By
FT SITE 285 291 similarity).
FT SITE 148 178 Beta-galactoside binding 2 (By
FT VARSPLIC 148 178 Missing (in isoform Short).
FT SEQUENCE 353 AA; 40036 MW; B54036F6E280C531 CRC64;
SQ SEQUENCE 353 AA; 40036 MW; B54036F6E280C531 CRC64;

Query Match 15.0%; Score 102; DB 1; Length 353;
Best Local Similarity 24.4%; Pred. No. 0.019;
Matches 32; Conservative 27; Mismatches 68; Indels 4; Gaps 2;

Qy 2 SPRLEVPCHALPQGLSPGVIIVRGLVLPQPKHTVSLR---DOAAHAPVTLRASFAADR 58
Db 218 TPATYPTPTIPNGLYPSKSMISGNVLPDTRPHNLRCGGDIAPHNLPRFNENAVVR 277

Qy 59 TLAWISRWGQ-KKLISAPFLFPQRFVLLFQEGGLKALNGGLGATSMNQALRQL 117
Db 278 NTQINNSWGQERSLLGRMPFSRGSFVSFWIICGHCXKVAVNGQHMCEYVHRLKNLQDI 337

Qy 118 RELRISGSVOL 128
Db 338 NTLVAGDIQL 348

```

```

RESULT 17
Q6QZP2 PRELIMINARY; PRT; 323 AA.
AC Q6QZP2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Galectin 9.
GN Name=UAT;
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Bannasch D.L., Ryon J.R., Bannasch M.J., Schaible R.H., Breen M.,
RA Ling G.;
RT "Exclusion of galectin 9 as a candidate gene for hyperuricosuria in
RT the Dalmatian dog";
RL Anim. Genet. 35:326-328 (2004).
CC -1- SIMILARITY: Belongs to the galectin (galectin/S-lectin) family.
DR EMBL; AY521549; AA580311.1; -.
DR HSP; P17931; IA3K.
DR GO; GO:0005529; F:sugar binding; IEA.
DR InterPro; IPR008985; ConA_like_lect_gl.
DR Pfam; PF00337; Galectin.
DR SMART; SM00276; GLECT; 2.
DR PROSITE; PS00309; GALAPTIN; 2.
KW Galectin; Lectin.
SQ SEQUENCE 323 AA; 36026 MW; 2C4FA3644A0EB64F CRC64;

Query Match 14.5%; Score 99; DB 2; Length 323;
Best Local Similarity 25.2%; Pred. No. 0.036;
Matches 33; Conservative 27; Mismatches 67; Indels 4; Gaps 2;

Qy 2 SPRLEPCSHALPGQLSPGQVIVRGVLQEPKHTVSLR---DQAAHAPVTLRASPADR 58
Db 188 TPTVPMFFTSIPGGLYPSKSIIVSGTVLPCKARFHNLRSGNDIAFHLNPRFNENTVVR 247
Qy 59 TLAWISRWG-OKKLISAPFLFYQRFPEVLLIFQEGGLKALNGGIGATSMNQQALEQL 117
Db 248 NMQINNWSGSEERSLPRKMPFVQGSFVIMCEGHCFKVAVDGEHLFYYHRLKNLAI 307
Qy 118 RELRISGSVOL 128
Db 308 NNMEVAGDVQL 318

RESULT 18
Q6P7Q6 PRELIMINARY; PRT; 322 AA.
AC Q6P7Q6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Lactin, galactose binding, soluble 9.
GN Name=Lgals9;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=Pituitary gland;
RX MEDLINE=23389257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,

Query Match 14.4%; Score 98; DB 2; Length 322;
Best Local Similarity 23.7%; Pred. No. 0.047;
Matches 31; Conservative 27; Mismatches 69; Indels 4; Gaps 2;

Qy 2 SPRLEPCSHALPGQLSPGQVIVRGVLQEPKHTVSLR---DQAAHAPVTLRASPADR 58
Db 187 TPTVPMFFTSIPGGLYPSKSIIVSGTVLPCKARFHNLRSGNDIAFHLNPRFNENTVVR 246
Qy 59 TLAWISRWG-OKKLISAPFLFYQRFPEVLLIFQEGGLKALNGGIGATSMNQQALEQL 117
Db 247 NTQINNWSGSEERSLPRKMPFVQGSFVIMCEGHCFKVAVDGEHLFYYHRLKNLAI 306
Qy 118 RELRISGSVOL 128
Db 307 NTLEVAGDIQL 317

RESULT 19
LEG4 MOUSE
ID LEG4 MOUSE STANDARD; PRT; 326 AA.
AC Q8K419; O88353; Q91X74;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE Galectin-4 (lactose-binding lectin 4).
GN Name=Lgals4;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=BALE/c; TISSUE=Liver;
RA Maly P., Jenikova G., Cummings R.D.;
RT "Molecular cloning and tissue distribution of mouse galectin-4.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6; TISSUE=Colon;
RA Hokama A., Tanaka Y., Mizoguchi A.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.

```



TISSUE=Colon; PubMed=12477932; DOI=10.1073/pnas.2426038999;  
MEDLINE=22388257; Feingold E.A., Grouse L.H., Derge J.G.,  
Strauberg R.L., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,  
Klausner R.D., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
Altschul S.F., Jordan B., Moore T., Max S.I., Wang J., Haieh F.,  
Hopkins R.F., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
Diatchenko L., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,  
Stapleton M., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
Brownstein M.J., Ulan R., Peters G.J., Abramson R.D., Mullahy S.J.,  
Raba S.S., Loquellano N.A., McKernan K.J., Malek J.A., Gunaratne P.H.,  
Bozak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
Fahy J., Helton E., Kettner M., Madan A.C., Rodrigues S., Sanchez A.,  
Whiting M., Madan A.C., Young A.C., Shevchenko Y., Bouffard G.G.,  
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
Rodriguez A.C., Greenwood J., Schmutz J., Myers R.M.,  
Butterfield Y.S.N., Krzyzinski M.I., Skalska U., Smalhus D.E.,  
Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
"Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences.";  
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
[4]  
SEQUENCE OF 22-321 FROM N.A., AND TISSUE SPECIFICITY.  
STRAIN=129/SV; TISSUE=Colon;  
MEDLINE=98112847; PubMed=9446608; DOI=10.1074/jbc.273.5.2954;  
Gitt M.A., Colnot C., Poirier F., Nani K.J., Barondes S.H.,  
Lefler H.;  
"Galectin-4 and galectin-6 are two closely related lectins expressed  
in mouse gastrointestinal tract.";  
J. Biol. Chem. 273:2954-2960 (1998).  
-!- FUNCTION: Galectin that binds lactose and a related range of  
sugars (By similarity).  
-!- SUBUNIT: Monomer (By similarity).  
-!- TISSUE SPECIFICITY: Epithelial cells of the embryonic and adult  
gastrointestinal tract. Expressed at about equal levels in colon  
and small intestine but much less in stomach.  
-!- DOMAIN: Contains two homologous but distinct carbohydrate-binding  
domains.  
-!- SIMILARITY: Belongs to the galectin (galactin/s-lectin) family.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; AY044870; AAK97790.1; -  
EMBL; AF510729; AAM44060.1; -  
EMBL; BC011236; AAH11236.1; -  
EMBL; BC021632; AAH21632.1; -  
EMBL; BC030297; AAH30297.1; -  
EMBL; AF026795; AAC27245.1; -  
HSSP; P17931; 1A3K.  
MGD; MGI:107536; Lgal54.  
InterPro; IPR008985; ConA\_like\_lect\_g1.  
InterPro; IPR001079; Galectin.  
Pfam; PF00337; Gal-bind\_lectin; 2.  
SMART; SM00276; GLECT; 2.  
PROSITE; PS00309; GALAPTIN; 2.  
KW Galectin; Lectin; Repeat.  
FT DOMAIN 1 152 Galaptin 1.  
FT DOMAIN 153 181 Linker.  
FT DOMAIN 182 326 Galaptin 2.  
FT SITE 259 265 Beta-galactoside binding (By similarity).  
FT SITE N -> K (in Ref. 2 and 4).  
FT CONFLICT 160 160 P -> A (in Ref. 4).  
FT CONFLICT 208 209 VR -> LP (in Ref. 4).  
FT CONFLICT 326 AA; 36372 MW; 7F3DD9862A85B5 CRC64;  
SQ SEQUENCE 326 AA; 36372 MW; 14.4k; Score 98; DB 1; Length 326;  
Query Match

Best Local Similarity 28.1%; Pred. No. 0.047;  
Matches 39; Conservative 24; Mismatches 54; Indels 22; Gaps 6;  
QY 2 SPRLEPCSHALPQGLSPGVIIIVRGVLVQEPKHTVSLR-----DQAAH-----APVTL 51  
DB 193 NPR--VPYVGLQGLVTRTIIGVYLVPTARNFVFKVSGSDIALHLNPRIGDSV 250  
QY 52 RASPADTLAWISRWG--QKLLSAPFLFPQRFVFEVLLLFQEGGLKALNQGGLGATSM 109  
DB 251 RNSFMN-----GSWGAERKVAYNP--FGPGQFFDLISIRGMDRFKVFANGQHLDFESH 302  
QY 110 NQQALEOLRELIRSGSVOL 128  
DB 303 RFQAFQWVDILEINGDITL 321  
RESULT 20  
LEG9 RAT STANDARD; PRT; 354 AA.  
AC P97840; O08588; O35866;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 25-JAN-2005 (Rel. 46, Last annotation update)  
DE Galectin-9 (36 kDa beta-galactoside binding lectin) (Urate  
DE transporter/channel) (UAT).  
GN Name=Lgals9;  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORMS LONG AND SHORT).  
RC STRAIN=Sprague-Dawley; TISSUE=Kidney, and Small intestine;  
RX MEDLINE=97190351; PubMed=9038233; DOI=10.1074/jbc.272.9.6078;  
RA Wada J., Kanwar Y.S.;  
RT "Identification and characterization of galectin-9, a novel beta-  
RT galactoside-binding mammalian lectin.";  
RL J. Biol. Chem. 272:6078-6086(1997).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORM SHORT).  
RC STRAIN=Sprague-Dawley; TISSUE=Kidney;  
RX MEDLINE=97150769; PubMed=8995305; DOI=10.1074/jbc.272.1.617;  
RA Leal-Pinto E., Tao W., Rappaport J., Richardson M., Knorr B.A.,  
RA Abramson R.G.;  
RT "Molecular cloning and functional reconstitution of a urate  
RT transporter/channel.";  
RL J. Biol. Chem. 272:617-625(1997).  
CC -!- FUNCTION: Binds galactosides. May play a role in thymocyte-  
CC epithelial interactions relevant to the biology of the thymus (By  
CC similarity). May provide the molecular basis for urate flux across  
CC cell membranes, allowing urate that is formed during purine  
CC metabolism to efflux from cells and serving as an electrogenic  
CC transporter that plays an important role in renal and  
CC gastrointestinal urate excretion. Highly selective to the anion  
CC urate.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic. May also be secreted by a non-  
CC classical secretory pathway (By similarity).  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Comment=Additional isoforms seem to exist;  
CC Name=Long;  
CC IsoId=P97840-1; Sequence=Displayed;  
CC Name=Short;  
CC IsoId=P97840-2; Sequence=VSP\_003098;  
CC -!- TISSUE SPECIFICITY: The long form is expressed exclusively in the  
CC small intestine.  
CC -!- DOMAIN: Contains two homologous but distinct carbohydrate-binding  
CC domains.  
CC -!- SIMILARITY: Belongs to the galectin (galaptin/s-lectin) family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR ENBL; U59462; AAB51192.1; -.
DR ENBL; U72741; AAB68592.1; -.
DR ENBL; U67958; AAB48591.1; -.
DR HSSP; P17931; IA3K.
DR RGD; 3005; Lgals9.
DR InterPro; IPR008985; ConA_like Lec_gl.
DR InterPro; IPR001079; Galectin.
DR Pfam; PF00337; Gal-bind lectin; 2.
DR PROSITE; PS00309; GALACTIN; 2.
KW Alternative splicing; Galectin; Ion transport; Lectin; Repeat.
FT DOMAIN 1 147 Galactin 1.
FT DOMAIN 148 205 Linker.
FT DOMAIN 206 354 Galactin 2.
FT SITE 81 87 Beta-galactoside binding 1 (By
FT SITE 286 292 Beta-galactoside binding 2 (By
FT SITE 148 179 Beta-galactoside binding 2 (By
FT VARSPLIC 148 179 Missing (in isoform Short).
FT SEQUENCE 354 AA; 39946 MW; 6574F960B2EAF37C CRC64;
Query Match 14.4%; Score 98; DB 1; Length 354;
Best Local Similarity 23.7%; Pred. No. 0.052;
Matches 31; Conservative 27; Mismatches 69; Indels 4; Gaps 2;
QY 2 SPRLVPCSHALPGLSPGVIIVRGLVLOEPKHTVSLR---DQAAHAPVTLRASPADR 58
DB 219 TPAYTIFFTSIPNGFYPSKSNINISGVLPDAKRFHINLRCCGGDIAPHLPNRFKVVVR 278
QY 59 TLAWISRWG-OKKLISAPFLFYPPQRFVLLLFQEGGLKALNGQGLGATSMNQALEQL 117
DB 279 NTQINNSWGFERSLPGRMFNRQGSFVILCEGHCFKAVDQGHICEYHRLKNLPDI 338
QY 118 RELRISGSVOL 128
DB 339 NTLVAGDIQL 349
RESULT 21
Q7TPX9 PRELIMINARY; PRT; 165 AA.
AC Q7TPX9
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Egg;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.; Feingold E.A.; Grouse L.H.; Derge J.G.;
RA Klausner R.D.; Collins F.S.; Wagner L.; Shenmen C.M.; Schuler G.D.;
RA Altschul S.F.; Zeeberg B.; Buetow K.H.; Schaefer C.F.; Bhat N.K.;
RA Hopkins R.F.; Jordan H.; Moore T.; Max S.I.; Wang J.; Hale F.;
RA Diatchenko L.; Marusina K.; Farmer A.A.; Rubin G.M.; Hong L.;
RA Stapleton M.; Soares M.B.; Bonaldo M.F.; Casavant T.L.; Scheetz T.E.;
RA Brownstein M.J.; Usdin T.B.; Toshiyuki S.; Carninci P.; Prange C.;
RA Raha S.S.; Iqbalano N.A.; Peters G.J.; Abramson R.D.; Mullaby S.J.;
RA Boak S.A.; McEwan P.J.; McKernan K.J.; Malek J.A.; Gunaratne P.H.;
RA Richards S.; Worley K.C.; Hale S.; Garcia A.M.; Gay L.J.; Hulyk S.W.;
RA Villalón D.K.; Muzny D.M.; Sodergren E.J.; Lu X.; Gibbs R.A.;
RA Fahey J.; Helton E.; Kettman M.; Madan A.; Rodriguez S.; Sanchez A.;
RA Whiting M.; Madan A.; Young A.C.; Shevchenko Y.; Bouffard G.G.;
RA Blakesley R.W.; Touchman J.W.; Green E.D.; Dickson M.C.;

```

```

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzyzinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Egg;
RA Strausberg R.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
CC -|- SIMILARITY: Belongs to the galectin (galactin/S-lectin) family.
DR ENBL; BC052838; AAH52838.1; -.
DR GO; GO:0005529; F:sugar binding; IEA.
DR InterPro; IPR008985; ConA_like Lec_gl.
DR InterPro; IPR001079; Galectin.
DR Pfam; PF00337; Gal-bind lectin; 1.
DR SMART; SM00276; GLECT; 1.
KW Galectin; Hypothetical protein; Lectin.
SQ SEQUENCE 165 AA; 18270 MW; 351BDDAF65B51D26 CRC64;
Query Match 14.3%; Score 97.5; DB 2; Length 165;
Best Local Similarity 24.6%; Pred. No. 0.024;
Matches 30; Conservative 29; Mismatches 54; Indels 9; Gaps 3;
QY 16 GLSPGVIIVRGLVLOEPKHTVSL-----RDQAAHAPVTLRASPADRTL---AWISRWG 67
DB 40 GMPGKVLVWGVVDHNPKSFALSITCGSDSEDPADVAIEKVVFTNQVFRNSCSGES 99
QY 68 OKKLISAP-FLFYPPQRFVLLLFQEGGLKALNGQGLGATSMNQALEQLRELIRISGV 126
DB 100 DEENLAFPPFPVDPDPFRMEIFCQPCFVLDVGHHLFDYHRIQTLAISAIKISGDL 159
QY 127 QL 128
DB 160 QI 161
RESULT 22
Q7ZTB8 PRELIMINARY; PRT; 308 AA.
AC Q7ZTB8
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Galectin family xgalectin-IIib.
GN Name=xgalectin-IIib;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=22552404; PubMed=12538594; DOI=10.1074/jbc.M209008200;
RA Shoji H.; Nishi N.; Hiraehima M.; Nakamura T.;
RT "Characterization of the Xenopus galectin family. Three structurally
RT different types as in mammals and regulated expression during
RT embryogenesis.";
RL J. Biol. Chem. 278:12285-12293(2003).
CC -|- SIMILARITY: Belongs to the galectin (galactin/S-lectin) family.
DR ENBL; AB080017; BAC55883.1; -.
DR HSP; P1116; ISLT.
DR GO; GO:0005529; F:sugar binding; IEA.
DR InterPro; IPR008985; ConA_like Lec_gl.
DR InterPro; IPR001079; Galectin.
DR Pfam; PF00337; Gal-bind lectin; 2.
DR SMART; SM00276; GLECT; 2.
KW Galectin; Lectin.
SQ SEQUENCE 308 AA; 34808 MW; A1C22D04079048AE CRC64;
Query Match 14.2%; Score 97; DB 2; Length 308;
Best Local Similarity 24.6%; Pred. No. 0.057;

```

```
Matches 32; Conservative 26; Mismatches 60; Indels 12; Gaps 4;
Qy 10 SHALPO-----GLSPGVIIIRGLVLQEPKHTVSLRDOAAHAPVTLRASFAADRTLAWI 63
Db 175 SYAMPYQNIYGLFPFKTIIVIRGTVPANPKRPHNLKXHGTA-LHNPFRDERTIVRN 233
Qy 64 SR-----WG-QKKLSAPFLFYQRFVFFVLLLFQEGGLKALNGOGLGATSMNQALBQLR 118
Db 234 SHLNGSWGNEERNLPRGMCFAFGSFVIEIRCEQHAQKVVNVNGAQICEFHRRVHQFOQID 293
Qy 119 ELRISGSVOL 128
Db 294 TLQIDGDVVL 303

RESULT 23
LEG6_MOUSE
ID LEG6_MOUSE STANDARD; PRT; 301 AA.
AC O54891; O88352;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Galectin-6.
GN Name=Lgals6;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Sv;
RX MEDLINE=98112847; PubMed=9446608; DOI=10.1074/jbc.273.5.2954;
RA Gitt M.A., Colnot C., Poirier F., Nani K.J., Barondes S.H.,
RA Leffler H.;
RT "Galectin-4 and galectin-6 are two closely related lectins expressed
RT in mouse gastrointestinal tract.";
RL J. Biol. Chem. 273:2954-2960(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Sv;
RX MEDLINE=98112848; PubMed=9446609; DOI=10.1074/jbc.273.5.2961;
RA Gitt M.A., Xia Y.-R., Atchison R.E., Luisi A.J., Barondes S.H.,
RA Leffler H.;
RT "Sequence, structure, and chromosomal mapping of the mouse Lgals6
RT gene, encoding galectin-6.";
RL J. Biol. Chem. 273:2961-2970(1998).
CC -!- DOMAIN: Contains two homologous but distinct carbohydrate-binding
CC domains.
CC -!- SIMILARITY: Belongs to the galectin (galaptin/s-lectin) family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF026799; AAC04508.1; -.
DR EMBL; AF026796; AAC04508.1; JOINED.
DR EMBL; AF026797; AAC04508.1; JOINED.
DR EMBL; AF026798; AAC04508.1; JOINED.
DR EMBL; AF026794; AAC27244.1; -.
DR HSSP; PI7931; 1A3K.
DR MGD; MGI:107535; Lgals6.
DR InterPro; IPR008985; ConA_like_lect_1.
DR InterPro; IPR001079; Galectin.
DR Pfam; PF00337; Gal-bind_lectin; 2.
DR SMART; SM00276; GLECT; 2.
DR PROSITE; PS00309; GALAPTIN; 2.
KW Galectin; Lectin; Repeat.
FT DOMAIN 1 151 Galectin 1.
FT DOMAIN 152 160 Linker.
```

```
FT DOMAIN 160 301 Galaptin 2.
FT CONFLICT 154 154 A -> V (in Ref. 2; AAC27244).
SQ SEQUENCE 301 AA; 34112 MW; 9A4DD09944EDFAB9 CRC64;

Query Match 14.1%; Score 96; DB 1; Length 301;
Best Local Similarity 27.2%; Pred. No. 0.071;
Matches 37; Conservative 22; Mismatches 61; Indels 16; Gaps 3;

Qy 7 VPCSHALPOGLSPGVIIIRGLVLQEPKHTVSLR-----DQAAH-----APVTLRASFA 56
Db 171 LPVVGALQGGFTVVRTIIKGYVLPKTAKTFAINFRVGSSEDIALHINPRIGDCLVRNSYM 230
Qy 57 DRTLAWISRWGOKKLISAPFLFYQRFVFFVLLLFQEGGLKALNGOGLGATSMNQALBQLR 116
Db 231 N-----GSMGTEERMVAYNPFPGGQGFDSLRCGMDRFKVFANGIHLFNFSHRFOALRK 284
Qy 117 LRELISGSVOLYCVH 132
Db 285 INTLEINGDITLSYVH 300

RESULT 24
O8QGD9 PRELIMINARY; PRT; 332 AA.
ID O8QGD9;
AC O8QGD9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Galectin-3TMI isoform containing transmembrane spanning domain.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22013973; PubMed=11866849; DOI=10.1074/jbc.M109578200;
RA Gorski J.P., Liu F.T., Artigues A., Castagna L.F., Osadoy P.;
RT "New alternatively spliced form of galectin-3, a member of the beta-
RT galactoside-binding animal lectin family, contains a predicted
RT transmembrane-spanning domain and a leucine zipper motif.";
RL J. Biol. Chem. 277:18840-18848(2002).
CC -!- SIMILARITY: Belongs to the galectin (galaptin/s-lectin) family.
DR EMBL; AF479564; AAL91920.1; -.
DR HSSP; PI7931; 1A3K.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005529; F:sugar binding; IEA.
DR InterPro; IPR008985; ConA_like_lect_1.
DR InterPro; IPR001079; Galectin.
DR Pfam; PF00337; Gal-bind_lectin; 1.
DR SMART; SM00276; GLECT; 1.
DR PROSITE; PS00309; GALAPTIN; UNKNOWN_1.
KW Galectin; Lectin; Transmembrane.
SQ SEQUENCE 332 AA; 35484 MW; CF3203BC99418CB1 CRC64;

Query Match 14.1%; Score 96; DB 2; Length 332;
Best Local Similarity 23.2%; Pred. No. 0.08;
Matches 33; Conservative 27; Mismatches 62; Indels 20; Gaps 3;

Qy 3 PRLEVPCSHALPOGLSPGVIIIRGLVLQEPKHTVSLR-----DQAAH-----APV 49
Db 193 PLQKVPYDPLPLPAGLMPRLITITGVNSNPNRFSLDKRGQDIAFHFNPRFKEDHKRVI 252
Qy 50 TLRASFAADRTLAWISRWGOKKLISAPFLFYQRFVFFVLLLFQEGGLKALNGOGLGATSM 109
Db 253 VCNEMFQ-----NNWKEERTAPRPFPGTGTGKLVQCEGDHFKVAVNDHALHQLQNF 305
Qy 110 NQQALEQLRELISGSVOLYCV 131
Db 306 REKKLNEITKLCIAGDIITLSV 327

RESULT 25
```





```

DR EMBL; BC054324; AAH54324.1; -.
DR HSP; P17931; 1A3K.
DR GO; GO:0005529; F:sugar binding; IEA.
DR InterPro; IPR008985; ConA like lec_gl.
DR InterPro; IPR001079; Galectin.
DR Pfam; PF00337; Gal-bind lectin; 1.
DR SMART; SM00276; GLECT; 1.
KW Galectin; Lectin.
FT NON TER
SQ SEQUENCE 204 AA; 22637 MW; 82532249A1339BE1 CRC64;

Query Match 13.8%; Score 94; DB 2; Length 204;
Best Local Similarity 24.3%; Pred. No. 0.074;
Matches 37; Conservative 25; Mismatches 62; Indels 28; Gaps 6;

QY 1 MSPRLEVPCSH-----ALPQGLSPQVIVRGLVLPQEPKPTVSLR-----DQAAH-- 46
DB 60 LPMWGPPIHPILPFKAMIPGGMIPKRTVIMKGLVNSNAKFIQISKVGYTNDIALHIN 119
QY 47 -----APVTLRASPADRTLAWISRWG--OKKLIISAPFLFYQRFEFVLLLFQEGGLKALN 100
DB 120 PRLNKNTLIIRNSFINGT-----WGEEDKDVKNP--FHQGEHFDISIRSGEKQKYVYN 171
QY 101 GQGLGATSMNQQAELRELIRISGSVOLYCVH 132
DB 172 GYHCFNYPHRLTNLQQVDTLEADGDIKLCFVH 203

RESULT 29
Q7ZTB6 PRELIMINARY; PRT; 319 AA.
AC Q7ZTB6;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Galectin family xgalectin-Via.
GN Xname=xgalectin-Via;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22552404; PubMed=12538594; DOI=10.1074/jbc.M209008200;
RA Shoji H., Nishi N., Hirashima M., Nakamura T.;
RT "Characterization of the Xenopus galectin family. Three structurally
RT different types as in mammals and regulated expression during
RT embryogenesis.";
RL J. Biol. Chem. 278:12285-12293 (2003).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391 (2002).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaney S.J.,
RA Boeak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

```

```

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skaleka U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Klein S., Gerhard D.S.;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the galectin (galactin/S-lectin) family.
DR EMBL; AB080019; BAC55885.1; -.
DR EMBL; BC077627; AAH77627.1; -.
DR HSP; P17931; 1A3K.
DR GO; GO:0005529; F:sugar binding; IEA.
DR InterPro; IPR008985; ConA like lec_gl.
DR InterPro; IPR001079; Galectin.
DR Pfam; PF00337; Gal-bind lectin; 2.
DR SMART; SM00276; GLECT; 2.
KW Galectin; Lectin.
SQ SEQUENCE 319 AA; 35369 MW; 2C12918A64B1BEBF CRC64;

Query Match 13.8%; Score 94; DB 2; Length 319;
Best Local Similarity 24.3%; Pred. No. 0.13;
Matches 37; Conservative 25; Mismatches 62; Indels 28; Gaps 6;

QY 1 MSPRLEVPCSH-----ALPQGLSPQVIVRGLVLPQEPKPTVSLR-----DQAAH-- 46
DB 175 LPMWGPPIHPILPFKAMIPGGMIPKRTVIMKGLVNSNAKFIQISKVGYTNDIALHIN 234
QY 47 -----APVTLRASPADRTLAWISRWG--OKKLIISAPFLFYQRFEFVLLLFQEGGLKALN 100
DB 235 PRLNKNTLIIRNSFINGT-----WGEEDKDVKNP--FHQGEHFDISIRSGEKQKYVYN 286
QY 101 GQGLGATSMNQQAELRELIRISGSVOLYCVH 132
DB 287 GYHCFNYPHRLTNLQQVDTLEADGDIKLCFVH 318

RESULT 30
Q8IGN9 PRELIMINARY; PRT; 218 AA.
AC Q8IGN9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative salivary galectin.
OS Anopheles stephensi (Indo-Pakistan malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=30069;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary glands;
RX MEDLINE=22710796; PubMed=12826099; DOI=10.1016/S0965-1748(03)00067-5;
RA Valenzuela J.G., Francischetti I.M.B., Pham V.M., Garfield M.K.,
RA Ribeiro J.M.C.;
RT "Exploring the salivary gland transcriptome and proteome of the
RT Anopheles stephensi mosquito.";
RL Insect Biochem. Mol. Biol. 33:717-732 (2003).
CC -!- SIMILARITY: Belongs to the galectin (galactin/S-lectin) family.
DR EMBL; AY162251; AAC06842.1; -.
DR HSP; P47929; 1BKZ.
DR GO; GO:0005529; F:sugar binding; IEA.
DR InterPro; IPR008985; ConA like lec_gl.
DR InterPro; IPR001079; Galectin.
DR Pfam; PF00337; Gal-bind lectin; 1.
DR SMART; SM00276; GLECT; 1.

```



**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2005, 08:36:15 ; Search time 172 Seconds

(without alignments)  
1315.435 Million cell updates/sec

Title: US-10-816-042-18

Perfect score: 3103

Sequence: 1 DAHKSEVAHRFKDLGENPK.....TCFAEKGKLVAAQAALGL 585

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

1: Geneseqp16Dec04:\*

2: Geneseqp1980s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003Bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	3103	100.0	585	1 AAR90388	Aap90388 Mature hu
2	3103	100.0	585	2 AAR05318	Aar05318 Human ser
3	3103	100.0	585	2 AAR08457	Aar08457 Human ser
4	3103	100.0	585	2 AAR80301	Aar80301 Human ser
5	3103	100.0	585	2 AAO20111	Aao20111 HSA prote
6	3103	100.0	585	3 AAY84873	Aay84873 Amino aci
7	3103	100.0	585	3 AAY83946	Aay83946 Yeast cod
8	3103	100.0	585	4 AAM52567	Aam52567 Mature hu
9	3103	100.0	585	4 AAE12403	Aae12403 Human alb
10	3103	100.0	585	4 AAE13129	Aae13129 Human alb
11	3103	100.0	585	4 AAE13399	Aae13399 Human alb
12	3103	100.0	585	4 ABB79006	Abb79006 Human mat
13	3103	100.0	585	4 AAE08578	Aae08578 Human ser
14	3103	100.0	585	5 AAU75220	Aau75220 Mature fo
15	3103	100.0	585	5 ABJ00986	Abj00986 B lymphoc
16	3103	100.0	585	5 ABG63321	Abg63321 Human ser
17	3103	100.0	585	5 ABG33847	Abg33847 Human B L
18	3103	100.0	585	5 ABG71291	Abg71291 Glycosyla
19	3103	100.0	585	6 ABR55695	Abr55695 Human alb
20	3103	100.0	585	7 ABR42606	Abr42606 Human ser
21	3103	100.0	585	7 ADC16767	Adc16767 Human ser
22	3103	100.0	585	7 ADD06469	Add06469 Human ser
23	3103	100.0	585	7 ADD68016	Add68016 Mature fo
24	3103	100.0	585	7 ADF15951	Adf15951 Human alb
25	3103	100.0	585	7 ADH21530	Adh21530 Mature hu

26	3103	100.0	585	7 ADK40322	Adk40322 Mature hu
27	3103	100.0	585	8 ABM79785	Abm79785 Human ser
28	3103	100.0	585	8 ADJ50489	Adj50489 Human ser
29	3103	100.0	585	8 ADJ64278	Adj64278 Human alb
30	3103	100.0	585	8 ADL76536	Adl76536 Albumin f
31	3103	100.0	585	8 ADP82617	Adp82617 Human alb
32	3103	100.0	585	8 ADP82618	Adp82618 Human ace
33	3103	100.0	585	8 ADR88124	Adr88124 Human alb
34	3103	100.0	609	3 AAB36542	Aab36542 Recombina
35	3103	100.0	609	3 AAY78147	Aay78147 Pre human
36	3103	100.0	609	3 AAB36549	Aab36549 Recombina
37	3103	100.0	609	6 ABUS7252	Abus7252 Human ser
38	3103	100.0	609	6 ABUS7253	Abus7253 Human ser
39	3103	100.0	609	7 ADD06471	Add06471 Human ser
40	3103	100.0	609	7 ADH21582	Adh21582 Human alb
41	3103	100.0	609	7 ADH21582	Adh21582 Human ser
42	3103	100.0	609	7 ADH21582	Adh21582 Human ser
43	3103	100.0	609	8 ADN16134	Adn16134 Protein s
44	3103	100.0	609	8 ADE77205	Ade77205 Human pro
45	3103	100.0	609	8 ADL13248	Adl13248 Human ste
46	3103	100.0	609	8 ADP55998	Adp55998 Human PRO
47	3103	100.0	610	2 AAR39510	Aar39510 Chimeric
48	3103	100.0	616	6 AAE30916	Aae30916 Val8-GLP-
49	3103	100.0	619	7 ADF14974	Adf14974 Human alb
50	3103	100.0	619	7 ADF14973	Adf14973 Human alb
51	3103	100.0	621	7 ADF16250	Adf16250 Human alb
52	3103	100.0	621	7 ADF16247	Adf16247 Human alb
53	3103	100.0	623	7 ADF16461	Adf16461 Human alb
54	3103	100.0	623	7 ADF16466	Adf16466 Human alb
55	3103	100.0	623	7 ADF16492	Adf16492 Human alb
56	3103	100.0	623	7 ADH21791	Adh21791 Human alb
57	3103	100.0	623	7 ADH21776	Adh21776 Human alb
58	3103	100.0	624	6 AAE30919	Aae30919 Human ser
59	3103	100.0	631	6 AAE30917	Aae30917 Val8-GLP-
60	3103	100.0	634	7 ADF16439	Adf16439 Human alb
61	3103	100.0	634	7 ADF16441	Adf16441 Human alb
62	3103	100.0	635	7 ADF16547	Adf16547 Human alb
63	3103	100.0	635	7 ADF15012	Adf15012 Human alb
64	3103	100.0	635	7 ADF15011	Adf15011 Human alb
65	3103	100.0	635	7 ADF15005	Adf15005 Human alb
66	3103	100.0	635	7 ADF15013	Adf15013 Human alb
67	3103	100.0	635	7 ADH21296	Adh21296 Human alb
68	3103	100.0	635	7 ADH21297	Adh21297 Human alb
69	3103	100.0	635	7 ADH21294	Adh21294 Human alb
70	3103	100.0	635	7 ADH21295	Adh21295 Human alb
71	3103	100.0	636	7 ADF16440	Adf16440 Human alb
72	3103	100.0	636	7 ADF16463	Adf16463 Human alb
73	3103	100.0	636	7 ADF16442	Adf16442 Human alb
74	3103	100.0	636	7 ADF16497	Adf16497 Human alb
75	3103	100.0	637	7 ADF16468	Adf16468 Human alb
76	3103	100.0	637	7 ADF16467	Adf16467 Human alb
77	3103	100.0	637	7 ADH21781	Adh21781 Human alb
78	3103	100.0	637	7 ADH21782	Adh21782 Human alb
79	3103	100.0	638	7 ADF16203	Adf16203 Human alb
80	3103	100.0	638	7 ADF42052	Adf42052 Plasmid p
81	3103	100.0	638	8 ADR90044	Adr90044 pDB2300X1
82	3103	100.0	639	7 ADF15120	Adf15120 Human alb
83	3103	100.0	639	7 ADF16198	Adf16198 Human alb
84	3103	100.0	639	7 ADF15119	Adf15119 Human alb
85	3103	100.0	639	7 ADF16208	Adf16208 Human alb
86	3103	100.0	639	7 ADF15118	Adf15118 Human alb
87	3103	100.0	639	7 ADF15116	Adf15116 Human alb
88	3103	100.0	639	7 ADH21332	Adh21332 Human alb
89	3103	100.0	639	7 ADH21335	Adh21335 Human alb
90	3103	100.0	639	7 ADH21333	Adh21333 Human alb
91	3103	100.0	639	7 ADH21334	Adh21334 Human alb
92	3103	100.0	640	6 AAE30920	Aae30920 Exendin-4
93	3103	100.0	640	6 AAE30918	Aae30918 Gly8-Glu2
94	3103	100.0	640	7 ADF14992	Adf14992 Human alb
95	3103	100.0	640	7 ADF16505	Adf16505 Human alb
96	3103	100.0	640	7 ADF16506	Adf16506 Human alb
97	3103	100.0	640	7 ADF16506	Adf16506 Human alb
98	3103	100.0	640	7 ADF14984	Adf14984 Human alb

99 3103 100.0 640 7 ADF16530 Human alb  
100 3103 100.0 640 7 ADH21818 Human alb

ALIGNMENTS

RESULT 1  
AAP90388  
ID AAP90388 standard; protein; 585 AA.

AC AAP90388;  
DT 24-OCT-2003 (revised)  
DT 25-MAR-2003 (revised)  
DT 01-NOV-1989 (first entry)  
XX Mature human serum albumin polypeptide.

XX Human serum albumin; mature protein; new polypeptides; plasma expanders.  
XX Homo sapiens; (Human).

XX EP322094-A.  
XX 28-JUN-1989.  
XX 25-OCT-1988; 88EP-00310000.  
XX 30-OCT-1987; 87GB-00025529.  
XX (DELZ ) DELTA BIOTECHNOLOGY LTD.

XX Ballance DJ, Hinchliffe E, Geisow MJ, Senior PJ;  
WPI; 1989-186464/26.  
DR N-PSDB; AAN90128.

XX New N-terminal fragments of human serum albumin - esp. useful as blood  
PT plasma expanders.

XX Disclosure; Fig 2; 20pp; English.  
XX Mature protein of human serum albumin (see corresp. AAN90128). Used to  
CC make new N-terminal fragments which are used as plasma expanders, or as  
CC substitutes for HSA or BSA, in tissue culture media. (Updated on 25-MAR-  
CC 2003 to correct PA field.) (Updated on 24-OCT-2003 to standardise OS  
CC field)

XX Sequence 585 AA;  
Query Match 100.0%; Score 3103; DB 1; Length 585;  
Best Local Similarity 100.0%; Pred. No. 1.9e-254;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRPKDLGEENFKALVLIATFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
DB 1 DAHSEVAHRPKDLGEENFKALVLIATFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
QY 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPMLPRLVRPEV 120  
DB 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPMLPRLVRPEV 120  
QY 121 DVMCTAFHDNETHFKKLYEYIARHPYFAPPELLFFAKRYKAAATECCQAADKAACLLP 180  
DB 121 DVMCTAFHDNETHFKKLYEYIARHPYFAPPELLFFAKRYKAAATECCQAADKAACLLP 180  
QY 181 KUDELURDEKASSAKORLKCASLQKGFGERAFKAWAVARLSQFPKAEFAEVSCLVTDLTJK 240  
DB 181 KUDELURDEKASSAKORLKCASLQKGFGERAFKAWAVARLSQFPKAEFAEVSCLVTDLTJK 240  
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSSISSKLKECCCKPILLESKSHCIAEVENDEMPA 300

DB 241 VHTCCCHGDLLECCADRADLAKYICENQDSSISSKLKECCCKPILLESKSHCIAEVENDEMPA 300  
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFILGMFLYEYARRHPDYSVLLARLAKTYETTLEKC 360  
DB 301 DLPSLAADFVESKDVCKNYAEAKDVFILGMFLYEYARRHPDYSVLLARLAKTYETTLEKC 360  
QY 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELFQELGGEYKFKQNALLVRYTKKVPQVST 420  
DB 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELFQELGGEYKFKQNALLVRYTKKVPQVST 420  
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCABDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCABDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480  
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVXHKPKAT 540  
DB 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVXHKPKAT 540  
QY 541 KEQLKAVNDDFAAFVEKCKCKADDDKTCFAEEGKKLVAASQAALGL 585  
DB 541 KEQLKAVNDDFAAFVEKCKCKADDDKTCFAEEGKKLVAASQAALGL 585

RESULT 2  
AAR05318  
ID AAR05318 standard; protein; 585 AA.

XX AAR05318;  
XX 08-OCT-1990 (first entry)  
XX Human serum albumin gene product.  
XX Human serum albumin; HSA-A; yeast; ds.  
XX Homo sapiens.  
XX JP02117384-A.  
XX 01-MAY-1990.

XX 26-OCT-1988; 88JP-00268302.  
XX 26-OCT-1988; 88JP-00268302.  
XX (TOFU ) TOA NENRYO KOGYO KK.  
XX WPI; 1990-176228/23.  
XX N-PSDB; AAQ04719.

XX Human serum albumin prepn. by yeast host - by culturing transformed  
PT plasmid yeast to produce serum, and removing it.

XX Disclosure; Page ?; -pp; Japanese.

CC Mature HSA-A may be produced using the sequence incorporated into a  
CC plasmid vector with suitable controllers, and transferred to a yeast  
CC expression system

XX Sequence 585 AA;

Query Match 100.0%; Score 3103; DB 2; Length 585;  
Best Local Similarity 100.0%; Pred. No. 1.9e-254;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRPKDLGEENFKALVLIATFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
DB 1 DAHSEVAHRPKDLGEENFKALVLIATFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
QY 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPMLPRLVRPEV 120  
DB 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPMLPRLVRPEV 120

```

QY 121 DVMCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
Db 121 DVMCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
QY 181 KLDELDRDEGKASSAKQRLKASLQKFGGERAFKAWAVARLSQRPFPKAEFAEVSCLVTDLTQ 240
Db 181 KLDELDRDEGKASSAKQRLKASLQKFGGERAFKAWAVARLSQRPFPKAEFAEVSCLVTDLTQ 240
QY 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKLKECEKPLLEKSHCIAEVENDEMPA 300
Db 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKLKECEKPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLVEYARRHPDYSVLLLLRLAKTYETTLK 360
Db 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLVEYARRHPDYSVLLLLRLAKTYETTLK 360
QY 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCLEFQELGEYKFNALLVRYTKVPQVST 420
Db 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCLEFQELGEYKFNALLVRYTKVPQVST 420
QY 421 PTLVEVSRNLGKVGSKCKKHPSAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 421 PTLVEVSRNLGKVGSKCKKHPSAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKQTALVELVKHKPKAT 540
Db 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKQTALVELVKHKPKAT 540
QY 541 KEQLKAVMDDDFAAFVEKCKCKADDKETCFABEGKLVAAASQAALGL 585
Db 541 KEQLKAVMDDDFAAFVEKCKCKADDKETCFABEGKLVAAASQAALGL 585

```

## RESULT 3

AAR08457  
ID AAR08457 standard; protein; 585 AA.

```

XX AAR08457;
XX AC
XX DT 25-MAR-2003 (revised)
XX DT 16-APR-1991 (first entry)
XX DE Human serum albumin.
XX KW HSA; folding; ss.
XX OS Homo sapiens.
XX FH Key
XX FT Region 1..303
XX FT /label= B
XX FT Region 123..585
XX FT /label= C
XX FT Region 123..303
XX FT /label= A
XX PN JP02227079-A.
XX PD
XX PF 10-SEP-1990.
XX PF 25-AUG-1989; 89JP-00217540.
XX PR 06-OCT-1988; 88JP-00250926.
XX PR (TOFU ) TONEN CORP.
XX PA
XX PX WPI; 1990-317325/42.
XX DR N-PSDB; AAQ06099.
XX XX
XX PT New human serum albumin fragments - used to bond to medicines and for
XX PT stable folding of protein(s).
XX PS Claim 1; Fig 8; 24pp; Japanese.

```

XX Fragments A-C of HSA are expressed as fusion proteins with the signal peptide of *E. coli* alkaline phosphatase. The fragments are selected for their specific properties. The C-terminal truncated fragment, B, does not bind long-chain fatty acids but does bind to various medicines at the central region. The N-terminal truncated fragment, C, has good stability in protein folding. The central segment, A, has characteristics of both B and C. See also AAQ06096-Q06098. (Updated on 25-MAR-2003 to correct PD field.) (Updated on 25-MAR-2003 to correct PF field.) (Updated on 25-MAR-2003 to correct PR field.)

XX Sequence 585 AA;

Query Match 100.0%; Score 3103; DB 2; Length 585;  
Best Local Similarity 100.0%; Pred. No. 1.9e-254; Indels 0; Gaps 0;  
Matches 585; Conservative 0; Mismatches 0;

```

QY 1 DAHKSEVAHRPKDLGEENFKALVLIATFAQYLOQCFFEDHVKLVNEVTEFAKTCVADESAB 60
Db 1 DAHKSEVAHRPKDLGEENFKALVLIATFAQYLOQCFFEDHVKLVNEVTEFAKTCVADESAB 60
QY 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 120
Db 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 120
QY 121 DVMCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
Db 121 DVMCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
QY 181 KLDELDRDEGKASSAKQRLKASLQKFGGERAFKAWAVARLSQRPFPKAEFAEVSCLVTDLTQ 240
Db 181 KLDELDRDEGKASSAKQRLKASLQKFGGERAFKAWAVARLSQRPFPKAEFAEVSCLVTDLTQ 240
QY 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKLKECEKPLLEKSHCIAEVENDEMPA 300
Db 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKLKECEKPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLVEYARRHPDYSVLLLLRLAKTYETTLK 360
Db 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLVEYARRHPDYSVLLLLRLAKTYETTLK 360
QY 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCLEFQELGEYKFNALLVRYTKVPQVST 420
Db 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCLEFQELGEYKFNALLVRYTKVPQVST 420
QY 421 PTLVEVSRNLGKVGSKCKKHPSAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 421 PTLVEVSRNLGKVGSKCKKHPSAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKQTALVELVKHKPKAT 540
Db 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKQTALVELVKHKPKAT 540
QY 541 KEQLKAVMDDDFAAFVEKCKCKADDKETCFABEGKLVAAASQAALGL 585
Db 541 KEQLKAVMDDDFAAFVEKCKCKADDKETCFABEGKLVAAASQAALGL 585

```

## RESULT 4

AAR0301  
ID AAR0301 standard; protein; 585 AA.

```

XX AAR0301;
XX AC
XX DT 25-MAR-2003 (revised)
XX DT 17-JAN-1996 (first entry)
XX DE Human serum albumin.
XX KW Serum albumin; HSA; aspartyl protease-3; Yap3p; Saccharomyces cerevisiae.
XX OS Homo sapiens.
XX XX

```

PN WO9523857-A1.  
XX  
PD 08-SEP-1995.  
XX  
PF 01-MAR-1995; 95WO-GB000434.  
XX  
PR 05-MAR-1994; 94GB-00004270.  
XX  
PA (DELZ ) DELTA BIOTECHNOLOGY LTD.  
XX  
PI KerryWilliams SM, Gilbert SC;  
XX  
DR WPI; 1995-320572/41.  
XX  
DR N-PSDB; AAQ98695.  
XX  
PT Yeast with reduced levels of aspartyl protease 3 proteolytic activity -  
PT used to secrete human albumin without prodn. of the 45 kD fragment.  
XX  
XX  
PS Example 1; Page 26-28; 50pp; English.  
XX  
CC The cDNA given in AAQ98695, which encodes HSA (AAR80301), was subjected  
CC to site-directed mutagenesis to investigate the role of endoproteases in  
CC the generation of a 45 kDa albumin fragment obtd. when the cDNA is  
CC expressed in S. cerevisiae. Mutations were: R410A, L407A, L408V, V409A;  
CC and R410A, K413Q, K414Q. The latter set of mutations, especially,  
CC improved stability of HSA to yeast Yap3p proteolytic cleavage, allowing  
CC increased prodn. of recombinant HSA. (Updated on 25-MAR-2003 to correct  
CC FI field.)  
XX  
SQ Sequence 585 AA;  
  
Query Match 100.0%; Score 3103; DB 2; Length 585;  
Best Local Similarity 100.0%; Pred. No. 1.9e-254; Indels 0; Gaps 0;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 DAHKSEVAHRFKDGLGEENFKALVLIAPAYLYQQCFPFEDHVKLVNEVTEFAKTCVADESAAE 60  
DB 1 DAHKSEVAHRFKDGLGEENFKALVLIAPAYLYQQCFPFEDHVKLVNEVTEFAKTCVADESAAE 60  
  
QY 61 NCDKSLHTLFGDKLTATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
DB 61 NCDKSLHTLFGDKLTATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
  
QY 121 DVMTCTAFHDNEETFLKKYLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
DB 121 DVMTCTAFHDNEETFLKKYLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
  
QY 181 KLDELRLDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQRPFAKAEFAVSKLVTDLTK 240  
DB 181 KLDELRLDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQRPFAKAEFAVSKLVTDLTK 240  
  
QY 241 VHTCECHGDLLECADRADLAKYICENQDSISSKLKECCPEPLLEKSHCIAEVENDEMPA 300  
DB 241 VHTCECHGDLLECADRADLAKYICENQDSISSKLKECCPEPLLEKSHCIAEVENDEMPA 300  
  
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYFYARRHPDYSVVLRLAKTYETTTLEKC 360  
DB 301 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYFYARRHPDYSVVLRLAKTYETTTLEKC 360  
  
QY 361 CAADPHECYAKVDFEPLVEEPQNLIKQNCLEFQELGEYKFNQALIVRYTKKVPQVST 420  
DB 361 CAADPHECYAKVDFEPLVEEPQNLIKQNCLEFQELGEYKFNQALIVRYTKKVPQVST 420  
  
QY 421 PTLVEVSNRLGKVGSKCKHPKAMPKCAEDYLSVLNQLCVLHKPTVSDRVTKCCTES 480  
DB 421 PTLVEVSNRLGKVGSKCKHPKAMPKCAEDYLSVLNQLCVLHKPTVSDRVTKCCTES 480  
  
QY 481 LVNRPFCFSALEVDSTYVPKFNARTFTFHADICTLSEKERQIKKQTALVELVHKPKAT 540  
DB 481 LVNRPFCFSALEVDSTYVPKFNARTFTFHADICTLSEKERQIKKQTALVELVHKPKAT 540  
  
QY 541 KEQLKAVNDDFAAFVEKCKADDKETCFABEGKKLVAASQAALGL 585  
DB 541 KEQLKAVNDDFAAFVEKCKADDKETCFABEGKKLVAASQAALGL 585

DB 541 KEQLKAVNDDFAAFVEKCKADDKETCFABEGKKLVAASQAALGL 585  
  
RESULT 5  
AAO20111  
ID AAO20111 standard; protein; 585 AA.  
XX  
XX AAO20111;  
AC AAO20111;  
XX  
DT 06-AUG-2002 (first entry)  
XX  
DE HSA protein sequence related to the growth hormone protein.  
XX  
KW Serum albumin-growth hormone fusion protein; growth hormone;  
KW Down's syndrome.  
XX  
OS Unidentified.  
XX  
PN KR99076789-A.  
XX  
PD 15-OCT-1999.  
XX  
PF 25-JUN-1998; 98KR-00704914.  
XX  
PR 30-DEC-1995; 95GB-00026733.  
PR 19-DEC-1996; 96WO-GB003164.  
XX  
PA (DELZ ) DELTA BIOTECHNOLOGY LTD.  
XX  
PI Ballance DJ;  
XX  
DR WPI; 1997-363680/33.  
DR N-PSDB; AAK99566.  
XX  
PT Serum albumin-growth hormone fusion protein - useful to treat growth  
PT hormone related diseases, e.g. Down's syndrome.  
XX  
PS Disclosure; Fig 6; 21pp; Korean.  
XX  
CC The invention relates to a serum albumin-growth hormone fusion protein -  
CC useful to treat growth hormone related diseases such as Down's syndrome.  
CC This sequence represents a HSA protein related to the serum albumin-  
CC growth hormone protein of the invention  
XX  
SQ Sequence 585 AA;  
  
Query Match 100.0%; Score 3103; DB 2; Length 585;  
Best Local Similarity 100.0%; Pred. No. 1.9e-254;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 DAHKSEVAHRFKDGLGEENFKALVLIAPAYLYQQCFPFEDHVKLVNEVTEFAKTCVADESAAE 60  
DB 1 DAHKSEVAHRFKDGLGEENFKALVLIAPAYLYQQCFPFEDHVKLVNEVTEFAKTCVADESAAE 60  
  
QY 61 NCDKSLHTLFGDKLTATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
DB 61 NCDKSLHTLFGDKLTATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
  
QY 121 DVMTCTAFHDNEETFLKKYLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
DB 121 DVMTCTAFHDNEETFLKKYLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
  
QY 181 KLDELRLDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQRPFAKAEFAVSKLVTDLTK 240  
DB 181 KLDELRLDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQRPFAKAEFAVSKLVTDLTK 240  
  
QY 241 VHTCECHGDLLECADRADLAKYICENQDSISSKLKECCPEPLLEKSHCIAEVENDEMPA 300  
DB 241 VHTCECHGDLLECADRADLAKYICENQDSISSKLKECCPEPLLEKSHCIAEVENDEMPA 300  
  
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYFYARRHPDYSVVLRLAKTYETTTLEKC 360  
DB 301 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYFYARRHPDYSVVLRLAKTYETTTLEKC 360

QY 361 CAADPHCYAKVDFEFKPLVBPQNLIKONCELFQQLGEYKFNALLVRYTKVPQVST 420  
DB 361 CAADPHCYAKVDFEFKPLVBPQNLIKONCELFQQLGEYKFNALLVRYTKVPQVST 420  
QY 421 PTLVEVSRNLGKVGSKCKHPKAKRMPKCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 421 PTLVEVSRNLGKVGSKCKHPKAKRMPKCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQOTALVELVGHKPKAT 540  
DB 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQOTALVELVGHKPKAT 540  
QY 541 KEQLKAVNMDDFAAVFEKCKCKADDDKTCFAEKGKLVAAASQAALGL 585  
DB 541 KEQLKAVNMDDFAAVFEKCKCKADDDKTCFAEKGKLVAAASQAALGL 585

RESULT 6  
AAV84873  
ID AAY84873 standard; protein; 585 AA.  
XX  
AC AAY84873;  
XX  
DT 08-AUG-2000 (first entry)  
XX  
DE Amino acid sequence of a human albumin protein.  
XX  
KW Human; albumin; ischemic state; serum protein; metal ion salt;  
KW peroperative ischemia; ischemia; myocardial infarction;  
KW progressive coronary artery disease.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 1  
FT /note= "optionally acetylated, and claimed under claim  
FT 56"  
XX  
PN WO200020840-A1.  
XX  
PD 13-APR-2000.  
XX  
PF 01-OCT-1999; 99WO-US022905.  
XX  
PR 02-OCT-1998; 98US-00165581.  
PR 02-OCT-1998; 98US-00165926.  
PR 02-OCT-1998; 98US-0102738P.  
PR 11-JAN-1999; 99US-0115392P.  
XX  
PA (ISCH-) ISCHEMIA TECHNOLOGIES INC.  
XX  
PI Bar-Or D, Lau E, Winkler JV;  
XX  
XX WPI; 2000-303843/26.  
XX

PT New method for the continuous detection of ischemic states comprises  
PT detecting and quantifying the existence of an alteration of the serum  
PT protein albumin.  
XX  
PS Disclosure; Page 97-100; 105pp; English.  
XX  
XX The present sequence represents human albumin protein. The specification  
XX describes a method for the continuous detection of ischemic states. The  
XX method comprises detecting and quantifying the existence of an alteration  
XX of the serum protein albumin. The method comprises contacting a  
XX biological sample containing albumin from the patient with an excess  
XX quantity of a metal ion salt, where the metal ion binds to the N-terminus  
XX of naturally occurring human albumin, to form a mixture containing bound  
XX metal ions and unbound metal ions, and then determining the amount of  
XX metal ions bound to the albumin N-terminus. The amount of bound metal  
XX ions is correlated to a known value to determine the occurrence or non-  
XX occurrence of an ischemic event. The methods are useful for detection of

CC ischemic states. The methods are also useful for distinguishing  
CC peroperative ischemia from ischemia caused by , amongst other things,  
CC myocardial infarctions and progressive coronary artery disease  
XX  
SQ Sequence 585 AA;  
Query Match 100.0%; Score 3103; DB 3; Length 585;  
Best Local Similarity 100.0%; Pred. No. 1.9e-254;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DAHKEVAHRPKDLGEENFKALVLIATAFYAQYLOQCPEFDHVKLVNVEVTEFAKTCVADESAE 60  
DB 1 DAHKEVAHRPKDLGEENFKALVLIATAFYAQYLOQCPEFDHVKLVNVEVTEFAKTCVADESAE 60  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECFLOHKDDNPRLVLRPEV 120  
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECFLOHKDDNPRLVLRPEV 120  
QY 121 DVMCTAFHNDNEETFLKKYLYEIARHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180  
DB 121 DVMCTAFHNDNEETFLKKYLYEIARHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180  
QY 181 KLDELURDEGKASSAKORLKASLOKFGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240  
DB 181 KLDELURDEGKASSAKORLKASLOKFGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240  
QY 241 VTECHCHGDLLECADDDRADLAKYICENQDSISSKLUCEKCEKPLLEKSHCIAEVENDEMPA 300  
DB 241 VTECHCHGDLLECADDDRADLAKYICENQDSISSKLUCEKCEKPLLEKSHCIAEVENDEMPA 300  
QY 301 DLPSLAADFVSKDKVCKNYAEAKDVLGMLFYEYARRHPDYSVLLLLLAKTYETTTLEKC 360  
DB 301 DLPSLAADFVSKDKVCKNYAEAKDVLGMLFYEYARRHPDYSVLLLLLAKTYETTTLEKC 360  
QY 361 CAADPHCYAKVDFEFKPLVBPQNLIKONCELFQQLGEYKFNALLVRYTKVPQVST 420  
DB 361 CAADPHCYAKVDFEFKPLVBPQNLIKONCELFQQLGEYKFNALLVRYTKVPQVST 420  
QY 421 PTLVEVSRNLGKVGSKCKHPKAKRMPKCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 421 PTLVEVSRNLGKVGSKCKHPKAKRMPKCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQOTALVELVGHKPKAT 540  
DB 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQOTALVELVGHKPKAT 540  
QY 541 KEQLKAVNMDDFAAVFEKCKCKADDDKTCFAEKGKLVAAASQAALGL 585  
DB 541 KEQLKAVNMDDFAAVFEKCKCKADDDKTCFAEKGKLVAAASQAALGL 585

RESULT 7  
AAV83946  
ID AAY83946 standard; protein; 585 AA.  
XX  
AC AAY83946;  
XX  
DT 28-JUL-2000 (first entry)  
XX  
DE Yeast codon-biased recombinant human serum albumin protein.  
XX  
KW Recombinant; human serum albumin; HSA; yeast codon bias; host cell;  
KW overlapping oligonucleotide; expression vector.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
PN CN1239103-A.  
XX  
PD 22-DEC-1999.  
XX  
PF 17-JUN-1998; 98CN-00102506.  
XX

```
PR 17-JUN-1998; 98CN-00102506.
XX (HAJ-) HALJI BIOENGINEERING CO LTD.
XX LI S, Lu D;
XX WPI; 2000-351198/31.
DR N-PSDB; AAA10091.
XX
XX Process for preparing recombinant human serum albumin comprising yeast
PT biased sex codons - uses a recombinant DNA technique.
XX Disclosure; Fig 1; 44pp; Chinese.
XX
XX The method relates to a method of recombinantly producing human serum
CC albumin (HSA) in yeast by altering the coding sequence of HSA to comprise
CC a yeast codon bias. The complete HSA gene (AAA10091) was generated as
CC three synthetic fragments (AAA10092-Al0094) joined by recombinant DNA
CC technology. Each HSA fragment was synthesised from overlapping
CC oligonucleotide fragments that were extended. This sequence represents
CC the complete sequence of the HSA gene encoded by the human gene with a yeast
CC codon bias. The invention also covers a recombinant expression vector,
CC yeast host cells carrying the recombinant expression vector and the
CC process for producing human serum albumin in the yeast host cell,
CC especially in secretory mode
XX
XX Sequence 585 AA;
SQ
Query Match 100.0%; Score 3103; DB 3; Length 585;
Best Local Similarity 100.0%; Pred. No. 1.9e-254;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHKEVAHRFKDLGEENFKALVLIAPAYLQOCFFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHKEVAHRFKDLGEENFKALVLIAPAYLQOCFFEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDESLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 61 NCDESLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
QY 121 DVMCTAFHDNEETFLKKLYEIARRHPYFYAPELFFAKRYKAAFTCCQAADKAACLPL 180
DB 121 DVMCTAFHDNEETFLKKLYEIARRHPYFYAPELFFAKRYKAAFTCCQAADKAACLPL 180
QY 181 KLDELDEGKASSAKORLKASLOKFGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 240
DB 181 KLDELDEGKASSAKORLKASLOKFGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 240
QY 241 VHTCECHGDLLECADRADLAKYICENODSISSSKLECEKPLLEKSHCIAEVENDEMPA 300
DB 241 VHTCECHGDLLECADRADLAKYICENODSISSSKLECEKPLLEKSHCIAEVENDEMPA 300
QY 301 DLPLSLAADFVESKDYCKNYAEAKDVFLGNFLYFYARRHPDYSVLLLLRLAKTYETTLK 360
DB 301 DLPLSLAADFVESKDYCKNYAEAKDVFLGNFLYFYARRHPDYSVLLLLRLAKTYETTLK 360
QY 361 CAAADPHCEYAKVDFEPFLVEEPQNLJKONCELFEOQLGEYKFQNALLVRYTKVPQVST 420
DB 361 CAAADPHCEYAKVDFEPFLVEEPQNLJKONCELFEOQLGEYKFQNALLVRYTKVPQVST 420
QY 421 PTLVEVSRLNLGVSKCKGHKPAKMPCAEDYLSVVLNQLCVLHBEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRLNLGVSKCKGHKPAKMPCAEDYLSVVLNQLCVLHBEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALVDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVGHKPKAT 540
DB 481 LVNRRPCFSALVDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVGHKPKAT 540
QY 541 KEQLKAVMDDFAAFVEKCKKADDKETCTFAEBGKGLVAASQAALGL 585
DB 541 KEQLKAVMDDFAAFVEKCKKADDKETCTFAEBGKGLVAASQAALGL 585
```

```
RESULT 8
AAM52567
ID AAM52567 standard; protein; 585 AA.
XX
XX AAM52567;
XX
XX 05-FEB-2002 (first entry)
XX
XX Mature human serum albumin.
XX
XX Human; serum albumin; HA; antiinflammatory; immunosuppressive; cardiant;
XX nontropic; neuroprotective; gene therapy; immune disorder; wound healing;
XX hyperproliferative disorder; renal disorder; cardiovascular disorder;
XX respiratory disorder; neurological disease; endocrine disorder;
XX reproductive system disorder; infectious disease;
XX gastrointestinal disorder.
XX
XX Homo sapiens.
XX
XX WO200179444-A2.
XX
XX 25-OCT-2001.
XX
XX 12-APR-2001; 2001WO-US012013.
XX
XX 12-APR-2000; 2000US-0229358P.
XX
XX 25-APR-2000; 2000US-0199384P.
XX
XX 21-DEC-2000; 2000US-0256931P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Haseltine WA;
XX
XX WPI; 2001-616755/71.
XX
XX N-PSDB; ABA03057.
XX
XX Albumin fusion proteins comprising a therapeutic protein and albumin,
XX useful in the treating immune system disorders (e.g. transplant
XX rejection), blood related disorders (e.g. myocardial infarction) and
XX hyperproliferative disorders.
XX
XX Claim 1; Fig 15; 606pp; English.
XX
XX The present invention relates to albumin fusion proteins, which comprise
XX a therapeutic protein and albumin. The present sequence is the protein
XX sequence for mature human serum albumin (HA), which was used to generate
XX the fusion proteins of the present invention. The albumin fusion proteins
XX are useful in the treatment, prevention, diagnosis, and/or detection of
XX diseases/disorders such as immune system disorders (e.g. transplant
XX rejection), blood related disorders (e.g. myocardial infarction),
XX hyperproliferative disorders (e.g. childhood acute myeloid leukemia),
XX renal disorders (e.g. glomerulonephritis), cardiovascular disorders (e.g.
XX arrhythmias), respiratory disorders (e.g. non-allergic rhinitis),
XX neurological diseases (e.g. Alzheimer's disease), endocrine disorders
XX (e.g. pheochromocytoma), reproductive system disorders (e.g. syphilis),
XX infectious diseases (e.g. measles), gastrointestinal disorders (e.g.
XX irritable bowel syndrome) and wound healing
XX
XX Sequence 585 AA;
SQ
Query Match 100.0%; Score 3103; DB 4; Length 585;
Best Local Similarity 100.0%; Pred. No. 1.9e-254;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHKEVAHRFKDLGEENFKALVLIAPAYLQOCFFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHKEVAHRFKDLGEENFKALVLIAPAYLQOCFFEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDESLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 61 NCDESLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
QY 121 DVMCTAFHDNEETFLKKLYEIARRHPYFYAPELFFAKRYKAAFTCCQAADKAACLPL 180
```

Db 121 DWCTAFHNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACTLP 180  
Qy 181 KLDELDRDEGKASSAKORLKASLOKFGGERAFKAWAVARLSORFPKAEFAEVSCLVTDLTk 240  
Db 181 KLDELDRDEGKASSAKORLKASLOKFGGERAFKAWAVARLSORFPKAEFAEVSCLVTDLTk 240  
Qy 241 VHTTECHGDLLECADDRLADLAKYICENQDSISSKLKECKEKLLEKSHCIAEVENDEMPA 300  
Db 241 VHTTECHGDLLECADDRLADLAKYICENQDSISSKLKECKEKLLEKSHCIAEVENDEMPA 300  
Qy 301 DLPSLAADPVEKDVCKNVAEAKDVFGLGMFLVEYARRHPDYSVLLLRLLAKTYETTLKc 360  
Db 301 DLPSLAADPVEKDVCKNVAEAKDVFGLGMFLVEYARRHPDYSVLLLRLLAKTYETTLKc 360  
Qy 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELFELQGEYKFONALLVRYTKKVPQVST 420  
Db 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELFELQGEYKFONALLVRYTKKVPQVST 420  
Qy 421 PTLVEVSRNLGVKSGCKCHPKAKMPCAEDYLSVVLNOLCVLHEKTPVSDRVTCKCTES 480  
Db 421 PTLVEVSRNLGVKSGCKCHPKAKMPCAEDYLSVVLNOLCVLHEKTPVSDRVTCKCTES 480  
Qy 481 LVNRRPCFSALVEDETYVPKEFNAETFTFHADICTLSEKEROIKKOTALVEILVKHKPKAT 540  
Db 481 LVNRRPCFSALVEDETYVPKEFNAETFTFHADICTLSEKEROIKKOTALVEILVKHKPKAT 540  
Qy 541 KEQLKAVMDDDFAAFVEKCKCKADDDKCTCFABEGKLVAAASQAALGL 585  
Db 541 KEQLKAVMDDDFAAFVEKCKCKADDDKCTCFABEGKLVAAASQAALGL 585

RESULT 9  
AAE12403  
ID AAE12403 standard; protein; 585 AA.  
AC AAE12403;  
XX  
DT 18-DEC-2001 (first entry)  
XX  
DE Human albumin (HA).  
XX  
KW Human; albumin; HA; immune system disorder; transplant rejection;  
KW blood related disorder; myocardial infarction; glomerulonephritis;  
KW hyperproliferative disorder; childhood acute myeloid leukaemia;  
KW renal cell carcinoma; cardiovascular disorder; vulnery; melanoma;  
KW arrhythmia; respiratory disorder; non-allergic rhinitis; antileukaemic;  
KW neurological disease; Alzheimer's disease; endocrine disorder; measles;  
KW pheochromocytoma; reproductive system disorder; neuroprotective; syphilis;  
KW infectious disease; gastrointestinal disorder; wound healing; nontropic;  
KW irritable bowel syndrome; HIV; human immunodeficiency virus infection;  
KW cytostatic; antiinflammatory; gene therapy; immunosuppressive; cardiant;  
KW antiarthritic; antirheumatic; renal disorder; antimicrobial.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Domain /label= Loop\_I  
FT Domain /label= Loop\_II  
FT Domain /label= Loop\_III  
FT Domain /label= Loop\_IV  
FT Domain /label= Loop\_V  
FT Domain /label= Loop\_VI  
FT Domain /label= Loop\_VII  
FT Domain /label= Loop\_VIII

FT Domain 439..447  
FT /label= Loop\_IX  
FT Domain 461..475  
FT /label= Loop\_X  
FT Domain 478..486  
FT /label= Loop\_XI  
FT Domain 560..566  
FT /label= Loop\_XII  
XX  
PN WO200179480-A1.  
XX  
PD 25-OCT-2001.  
XX  
PF 12-APR-2001; 2001WO-US011991.  
XX  
PR 12-APR-2000; 2000US-0229358P.  
PR 25-APR-2000; 2000US-0199384P.  
PR 21-DEC-2000; 2000US-0256931P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Haseltine WA;  
XX  
DR WPI: 2001-616756/71.  
DR N-PSDB; AAD20005.  
XX  
PT Albumin fusion proteins comprising a therapeutic protein and albumin,  
PT useful in the treating metastatic renal cell carcinoma, metastatic  
PT melanoma, malignant melanoma, renal cell carcinoma, HIV (human  
PT immunodeficiency virus) or infection.  
XX  
PS Claim 1; Fig 9; 394pp; English.  
XX  
CC The invention relates to human albumin (HA) fusion proteins and their  
CC corresponding nucleic acid sequences. Therapeutic proteins fused to  
CC albumin or its fragments have an extended shelf-life. The albumin fusion  
CC proteins are useful in the treatment, prevention, diagnosis, and/or  
CC detection of diseases, disorders such as immune system disorders (e.g.  
CC transplant rejection), blood related disorders (e.g. myocardial  
CC infarction), hyperproliferative disorders (e.g. childhood acute myeloid  
CC leukaemia, metastatic renal cell carcinoma, metastatic melanoma,  
CC malignant melanoma, renal cell carcinoma), renal disorders (e.g.  
CC glomerulonephritis), cardiovascular disorders (e.g. arrhythmias), diseases  
CC respiratory disorders (e.g. non-allergic rhinitis), neurological diseases  
CC (e.g. Alzheimer's disease), endocrine disorders (e.g. pheochromocytoma),  
CC reproductive system disorders (e.g. syphilis), infectious diseases (e.g.  
CC measles), gastrointestinal disorders (e.g. irritable bowel syndrome), HIV  
CC (human immunodeficiency virus) infection and wound healing. Nucleic acids  
CC encoding albumin fusion protein is used in gene therapy. The present  
CC sequence is human albumin  
SQ Sequence 585 AA;  
Query Match 100.0%; Score 3103; DB 4; Length 585;  
Best Local Similarity 100.0%; Pred. No. 1.9e-254; Mismatches 0; Indels 0; Gaps 0;  
Matches 585; Conservative 0;  
Qy 1 DAHKSEVAHRFKDLGEENFKALVLIATAQYLOQCPEFDHVKLVNEVTEFAKTCVADESAE 60  
Db 1 DAHKSEVAHRFKDLGEENFKALVLIATAQYLOQCPEFDHVKLVNEVTEFAKTCVADESAE 60  
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQSPERNECFLOHKDDNPRLVRPEV 120  
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQSPERNECFLOHKDDNPRLVRPEV 120  
Qy 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACTLP 180  
Db 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACTLP 180  
Qy 181 KLDELDRDEGKASSAKORLKASLOKFGGERAFKAWAVARLSORFPKAEFAEVSCLVTDLTk 240  
Db 181 KLDELDRDEGKASSAKORLKASLOKFGGERAFKAWAVARLSORFPKAEFAEVSCLVTDLTk 240



QY	241	VHTECHGDLLECADRADLAKYICENQDSISSKLKECEKPLLEKSHCIAEVENDEMPA	300
Db	241	VHTECHGDLLECADRADLAKYICENQDSISSKLKECEKPLLEKSHCIAEVENDEMPA	300
QY	301	DLPSLAADFVESKOVCKNYAEAKOVFLGMFLYFYARRHPDYSVLLLRSLAKTYETTLK	360
Db	301	DLPSLAADFVESKOVCKNYAEAKOVFLGMFLYFYARRHPDYSVLLLRSLAKTYETTLK	360
QY	361	CAAADPHCYAKVDFEFPLVEEPQNLIKONCELFQQLGEYKFNQALLVRYTKVPQVST	420
Db	361	CAAADPHCYAKVDFEFPLVEEPQNLIKONCELFQQLGEYKFNQALLVRYTKVPQVST	420
QY	421	PTLVEVSNLGVSKCKKHPKAMPKCAEDYLSVNLQCLVHLKTPVSDRVTKCCTES	480
Db	421	PTLVEVSNLGVSKCKKHPKAMPKCAEDYLSVNLQCLVHLKTPVSDRVTKCCTES	480
QY	481	LVNRRPCFSALEVDYTYVPKFNPAETFTFHADICTLSEKQRIKKQTALVELVKHKPKAT	540
Db	481	LVNRRPCFSALEVDYTYVPKFNPAETFTFHADICTLSEKQRIKKQTALVELVKHKPKAT	540
QY	541	KEQLKAVMDDFAAFVEKCKKADDDKTCFAEKGKLVAAASQAALGL	585
Db	541	KEQLKAVMDDFAAFVEKCKKADDDKTCFAEKGKLVAAASQAALGL	585
RESULT 10			
AAE13129			
ID	AAE13129 standard; protein; 585 AA.		
AC	AAE13129;		
XX			
DT	28-JAN-2002 (first entry)		
XX			
DE	Human albumin (HA).		
XX			
KW	Human; albumin; HA; fusion protein; therapeutic protein; vulnery;		
KW	immune system disorder; transplant rejection; blood related disorder;		
KW	myocardial infarction; hyperproliferative disorder; glomerulonephritis;		
KW	childhood acute myeloid leukemia; cardiovascular disorder; arrhythmia;		
KW	respiratory disorder; gene therapy; non-allergic rhinitis; neutropic;		
KW	neurological disease; Alzheimer's disease; reproductive system disorder;		
KW	endocrine disorder; pheochromocytoma; infectious disease; antiarthritic;		
KW	measles; gastrointestinal disorder; irritable bowel syndrome; syphilis;		
KW	wound healing; antiinflammatory; immunosuppressive; neuroprotective;		
KW	cardiant; cytostatic; antileukaemic; antirheumatic; antimicrobial;		
KW	renal disorder.		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	Domain	54..61	
FT		/label= Loop_I	
FT	Domain	76..89	
FT		/label= Loop_II	
FT	Domain	92..100	
FT		/label= Loop_III	
FT	Domain	170..176	
FT		/label= Loop_IV	
FT	Domain	247..252	
FT		/label= Loop_V	
FT	Domain	266..277	
FT		/label= Loop_VI	
FT	Domain	280..288	
FT		/label= Loop_VII	
FT	Domain	362..368	
FT		/label= Loop_VIII	
FT	Domain	439..447	
FT		/label= Loop_IX	
FT	Domain	461..475	
FT		/label= Loop_X	
FT	Domain	478..486	
FT		/label= Loop_XI	
FT	Domain	560..566	

FT		/label= Loop_XII	
XX			
PN	WO200179443-A2.		
XX			
PD	25-OCT-2001.		
XX			
PF	12-APR-2001; 2001WO-US011924.		
XX			
PR	12-APR-2000; 2000US-0229358P.		
PR	25-APR-2000; 2000US-0199384P.		
PR	21-DEC-2000; 2000US-0256931P.		
XX			
PA	(HUMA-) HUMAN GENOME SCI INC.		
XX			
PI	Rosen CA, Haseltine WA;		
XX			
DR	WPI; 2001-616754/71.		
DR	N-PSDB; AAD21638.		
XX			
PT	Albumin fusion proteins comprising a therapeutic protein and albumin,		
PT	useful in the treating immune system disorders (e.g. transplant		
PT	rejection), blood related disorders (e.g. myocardial infarction) and		
PT	hyperproliferative disorders.		
XX			
PS	Claim 1; Fig 9; 380pp; English.		
XX			
CC	The invention relates to albumin fusion proteins comprising therapeutic		
CC	protein and human albumin (HA). Therapeutic proteins fused to albumin have		
CC	an extended shelf-life. The albumin fusion proteins are useful in the		
CC	treatment, prevention, diagnosis and/or detection of diseases, disorders		
CC	such as immune system disorders (e.g. transplant rejection), blood		
CC	related disorders (e.g. myocardial infarction), hyperproliferative		
CC	disorders (e.g. childhood acute myeloid leukaemia), renal disorders (e.g.		
CC	glomerulonephritis), cardiovascular disorders (e.g. arrhythmias),		
CC	respiratory disorders (e.g. non-allergic rhinitis), neurological diseases		
CC	(e.g. Alzheimer's disease), endocrine disorders (e.g. pheochromocytoma),		
CC	reproductive system disorders (e.g. syphilis), infectious diseases (e.g.		
CC	measles), gastrointestinal disorders (e.g. irritable bowel syndrome) and		
CC	wound healing. Nucleic acids encoding albumin fusion protein is used in		
CC	gene therapy. The present sequence is human albumin (HA) protein		
XX			
SQ	Sequence 585 AA;		
	Query Match	100.0%; Score 3103; DB 4; Length 585;	
	Best Local Similarity	100.0%; Pred. No. 1.9e-254;	
	Matches 585; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	DAHKSEVAHRFKDLGEENFKALVLI	FAQYLQOCPPFEDHVKLVNEVTEFAKTCVADES 60
Db	1	DAHKSEVAHRFKDLGEENFKALVLI	FAQYLQOCPPFEDHVKLVNEVTEFAKTCVADES 60
QY	61	NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV	120
Db	61	NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV	120
QY	121	DVMCTAFHDNEETFLKXYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP	180
Db	121	DVMCTAFHDNEETFLKXYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP	180
QY	181	KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRFPAEFAEVS	KLVTDLTK 240
Db	181	KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRFPAEFAEVS	KLVTDLTK 240
QY	241	VHTECHGDLLECADRADLAKYICENQDSISSKLKECEKPLLEKSHCIAEVENDEMPA	300
Db	241	VHTECHGDLLECADRADLAKYICENQDSISSKLKECEKPLLEKSHCIAEVENDEMPA	300
QY	301	DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVVLLRLAKTYETTLK	360
Db	301	DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVVLLRLAKTYETTLK	360
QY	361	CAAADPHCYAKVDFEFKPLVEEPQNLIKONCELFQQLGEYKFNQALLVRYTKVPQVST	420



Db 361 CAADPHECYAKVDFEFKPLVBPQNLIKQNCLEFQOLGEYKFQNALLVRYTKVPQVST 420  
QY 421 PTLVEVSRNLGVSKCCRRHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480  
Db 421 PTLVEVSRNLGVSKCCRRHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480  
QY 481 LVNRRPCFSALEVDITYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKGHPKAT 540  
Db 481 LVNRRPCFSALEVDITYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKGHPKAT 540  
QY 541 KEQLKAVMDDDFAAFVEKCKCADDKETCFABEGKLVAAASQAALGL 585  
Db 541 KEQLKAVMDDDFAAFVEKCKCADDKETCFABEGKLVAAASQAALGL 585

RESULT 11  
AAE13399  
ID AAE13399 standard; protein; 585 AA.  
XX  
AC AAE13399;  
XX  
DT 12-FEB-2002 (first entry)  
XX  
DE Human albumin (HA) protein.  
XX  
KW Human; albumin; HA; fusion protein; immune system disorder; syphilis;  
KW transplant rejection; blood related disorder; myocardial infarction;  
KW hyperproliferative disorder; acute myeloid leukaemia; renal disorder;  
KW glomerulonephritis; cardiovascular disorder; arrhythmia; rhinitis;  
KW respiratory disorder; neurological disease; Alzheimer's disease;  
KW endocrine disorder; pheochromocytoma; reproductive system disorder;  
KW measles; gastrointestinal disorder; irritable bowel syndrome; HIV;  
KW human immunodeficiency virus; wound healing; renal cell carcinoma;  
KW melanoma; gene therapy.

OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Domain 54..61  
FT /label= Loop\_I  
FT Domain 76..89  
FT /label= Loop\_II  
FT Domain 92..100  
FT /label= Loop\_III  
FT Domain 170..176  
FT /label= Loop\_IV  
FT Domain 247..252  
FT /label= Loop\_V  
FT Domain 266..277  
FT /label= Loop\_VI  
FT Domain 280..288  
FT /label= Loop\_VII  
FT Domain 362..368  
FT /label= Loop\_VIII  
FT Domain 439..447  
FT /label= Loop\_IX  
FT Domain 461..475  
FT /label= Loop\_X  
FT Domain 478..486  
FT /label= Loop\_XI  
FT Domain 560..566  
FT /label= Loop\_XII  
XX  
PN WO200179258-A1.  
PD 25-OCT-2001.  
XX  
PF 12-APR-2001; 2001WO-US012008.  
XX  
PR 12-APR-2000; 2000US-0229358P.  
PR 25-APR-2000; 2000US-0199384P.  
PR 21-DEC-2000; 2000US-0256931P.  
XX

(HUMA-) HUMAN GENOME SCI INC.  
(PRIN-) PRINCIPIA PHARM CORP.  
Rosen CA, Sadeghi H, Prior CP, Turner AJ;  
WPI: 2001-602931/68.  
N-PSDB; AAD22287.  
Albumin fusion proteins comprising a therapeutic protein and albumin,  
useful in the treating metastatic renal cell carcinoma, metastatic  
melanoma, malignant melanoma, renal cell carcinoma, HIV (human  
immunodeficiency virus) or infection.  
Claim 1; Fig 9; 325pp; English.  
XX  
PS The invention relates to albumin fusion proteins comprising therapeutic  
protein and human albumin (HA). The albumin fusion proteins are useful in  
the treatment, prevention, diagnosis, and/or detection of diseases;  
disorders such as immune system disorders (transplant rejection); blood  
related disorders (myocardial infarction); hyperproliferative disorders  
(childhood acute myeloid leukaemia); renal disorder (glomerulonephritis);  
cardiovascular disorders (arrhythmias); respiratory disorders (non-  
allergic rhinitis); neurological diseases (Alzheimer's disease);  
endocrine disorders (pheochromocytoma); reproductive system disorders  
(syphilis); infectious diseases (measles); gastrointestinal disorders  
(irritable bowel syndrome) and wound healing. The albumin fusion proteins  
are also used in the treatment of metastatic renal cell carcinoma,  
metastatic melanoma, malignant melanoma and HIV (human immunodeficiency  
virus) infection. Nucleic acid encoding albumin fusion protein is useful  
in gene therapy. The present sequence is human albumin (HA) protein  
Sequence 585 AA;  
SQ  
Query Match 100.0%; Score 3103; DB 4; Length 585;  
Best Local Similarity 100.0%; Pred. No. 1.9e-254;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DAHSEVAHRFXDLGEENFKALVLIATAFYLOQCQPFEDHVKLVNEVTEFAKTCVADESAE 60  
Db 1 DAHSEVAHRFXDLGEENFKALVLIATAFYLOQCQPFEDHVKLVNEVTEFAKTCVADESAE 60  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQPERNECFLOHDDNPNLRLVRPEV 120  
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQPERNECFLOHDDNPNLRLVRPEV 120  
QY 121 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQQAADKAACLLP 180  
Db 121 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQQAADKAACLLP 180  
QY 181 KLDELRLDEGKASSAKORLKCASLQKFGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 240  
Db 181 KLDELRLDEGKASSAKORLKCASLQKFGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 240  
QY 241 VHTCECHGDLLECADRADLAKYICENQDSISSKLEKCECEKELLESKSHCIAEVENDEMPA 300  
Db 241 VHTCECHGDLLECADRADLAKYICENQDSISSKLEKCECEKELLESKSHCIAEVENDEMPA 300  
QY 301 DLPSLAADFVESKDYCKNVAEAKDVFLGMFLYEYARRHPDYSVLLRLRAKTYETTLK 360  
Db 301 DLPSLAADFVESKDYCKNVAEAKDVFLGMFLYEYARRHPDYSVLLRLRAKTYETTLK 360  
QY 361 CAADPHECYAKVDFEFKPLVBPQNLIKQNCLEFQOLGEYKFQNALLVRYTKVPQVST 420  
Db 361 CAADPHECYAKVDFEFKPLVBPQNLIKQNCLEFQOLGEYKFQNALLVRYTKVPQVST 420  
QY 421 PTLVEVSRNLGVSKCCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480  
Db 421 PTLVEVSRNLGVSKCCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480  
QY 481 LVNRRPCFSALEVDITYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKGHPKAT 540  
Db 481 LVNRRPCFSALEVDITYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKGHPKAT 540

Qy 541 KEQLKAVNDDFAAFVEKCKKADDETCFAEBGKGLVAASQAALGL 585  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||  
541 KEQLKAVNDDFAAFVEKCKKADDETCFAEBGKGLVAASQAALGL 585  
RESULT 12  
ABB79006  
ID ABB79006 standard; protein; 585 AA.  
XX  
AC ABB79006;  
XX  
DT 01-AUG-2002 (first entry)  
XX  
DE Human mature albumin protein SEQ ID NO:18.  
XX  
KW Human; growth hormone; hGH; albumin; human serum albumin; HSA;  
KW albumin fusion protein; cytosstatic; anorectic; immunosuppressive;  
KW antidiabetic; antirheumatic; antiarthritic; psoriatic; cancer;  
KW non-Hodgkin's lymphoma; obesity; transplant rejection; psoriasis;  
KW type I diabetes mellitus; rheumatoid arthritis.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Domain 1..194  
FT /label= 1  
FT Domain 1..105  
FT /label= subdomain  
FT Disulfide-bond 53..62  
FT Disulfide-bond 75..91  
FT Disulfide-bond 90..101  
FT Region 106..119  
FT /note= "flexible inter-subdomain linker region"  
FT Domain 120..194  
FT /label= subdomain  
FT Disulfide-bond 124..169  
FT Disulfide-bond 168..177  
FT Domain 195..387  
FT /label= 2  
FT Domain 195..291  
FT /label= subdomain  
FT Disulfide-bond 245..253  
FT Disulfide-bond 265..279  
FT Disulfide-bond 278..289  
FT Region 292..315  
FT /note= "flexible inter-subdomain linker region"  
FT Domain 316..387  
FT /label= subdomain  
FT Disulfide-bond 316..361  
FT Disulfide-bond 360..369  
FT Domain 388..585  
FT /label= 3  
FT Domain 388..491  
FT /label= subdomain  
FT Disulfide-bond 392..438  
FT Disulfide-bond 437..448  
FT Disulfide-bond 461..477  
FT Disulfide-bond 476..487  
FT Region 492..511  
FT /note= "flexible inter-subdomain linker region"  
FT Domain 512..585  
FT /label= subdomain  
FT Disulfide-bond 514..559  
FT Disulfide-bond 558..567  
XX  
WO200179442-A2.  
XX  
PD 25-OCT-2001.  
XX  
PF 12-APR-2001; 2001WO-US011850.  
XX  
PR 12-APR-2000; 2000US-0229358P.  
PR 25-APR-2000; 2000US-0199384P.

PR 21-DEC-2000; 2000US-0256931P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Haseltine WA;  
XX  
DR WPI; 2001-611723/70.  
DR N-PSDB; ABN87288.  
XX  
PT New albumin fusion proteins, useful for treating diseases and disorders  
PT such as cancer, comprise therapeutic protein fused to albumin.  
XX  
PS Claim 1; Fig 11; 413pp; English.  
XX  
CC The present invention describes an albumin fusion protein (I) comprising  
CC a therapeutic protein: X and (a fragment or variant of) albumin  
CC comprising the fully defined sequence in ABB79006 of 585 amino acids,  
CC (where the fragment or variant has albumin or therapeutic protein: X  
CC activity). (I) can have cytostatic, anorectic, immunosuppressive,  
CC antidiabetic, antirheumatic, antiarthritic and psoriatic activities.  
CC Albumin fusion proteins are stabilised therapeutic proteins e.g  
CC antibodies to CS, C242 and CD80 useful for treating various diseases and  
CC disorders such as non-Hodgkin's lymphoma, cancer, obesity, transplant  
CC rejection, type I diabetes mellitus, rheumatoid arthritis and psoriasis.  
CC Fusing albumin to therapeutic proteins stabilises the therapeutic  
CC protein, extends the shelf life and retains the in vitro or in vivo  
CC biological activity. It also reduces the need to formulate protein  
CC solutions with large excesses of carrier proteins to prevent loss of  
CC therapeutic proteins due to factors such as binding to the container. The  
CC fusion proteins are easily dispensed with a simple formulation requiring  
CC minimal post storage manipulation. The fusion of therapeutic proteins to  
CC albumin confers stability in aqueous or other solution. The present  
CC sequence represents the mature human albumin (HA) protein which is used  
CC in the exemplification of the present invention  
XX  
SQ Sequence 585 AA;  
Query Match 100.0%; Score 3103; DB 4; Length 585;  
Best Local Similarity 100.0%; Pred. No. 1.9e-254;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 DAHKSEVAHRFKDGLGEENFKALVLIAPAYQLQQCFEDHVKLVNVEVTFKTCVADESAAE 60  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||  
1 DAHKSEVAHRFKDGLGEENFKALVLIAPAYQLQQCFEDHVKLVNVEVTFKTCVADESAAE 60  
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPRLVRPEV 120  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||  
61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPRLVRPEV 120  
Qy 121 DVMTAFHDNEETFLKYLVEIARRHPYFYAPELFFAKRYKAAFTCCQAADKAACLLP 180  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||  
121 DVMTAFHDNEETFLKYLVEIARRHPYFYAPELFFAKRYKAAFTCCQAADKAACLLP 180  
Qy 181 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRFPAEFAEVSCLVTDLT 240  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||  
181 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRFPAEFAEVSCLVTDLT 240  
Qy 241 VHTECHGDLLECADDDRADLAKYICENQDSISSKLKECEKPLLEKSHCIAEVENDEMPA 300  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||  
241 VHTECHGDLLECADDDRADLAKYICENQDSISSKLKECEKPLLEKSHCIAEVENDEMPA 300  
Qy 301 DLPSLAADFVESKQVKNYAEAKDVFLGMFLYEYARRHPDYSVLLLLLAKTYETTLEKC 360  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||  
301 DLPSLAADFVESKQVKNYAEAKDVFLGMFLYEYARRHPDYSVLLLLLAKTYETTLEKC 360  
Qy 361 CAADPHECYAKVDFEFPKPLVEEPQNLIKQNCELFEQLGEYKFQNALVRYTKVPQVST 420  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||  
361 CAADPHECYAKVDFEFPKPLVEEPQNLIKQNCELFEQLGEYKFQNALVRYTKVPQVST 420  
Qy 421 PTLVEVSRNLGKVGSKCKGHPKAPCAEDVLSVVLNQLCVLHEKTPVSDRVTTCCTES 480  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||  
421 PTLVEVSRNLGKVGSKCKGHPKAPCAEDVLSVVLNQLCVLHEKTPVSDRVTTCCTES 480

QY 481 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVKKPKAT 540  
DB LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVKKPKAT 540  
QY 541 KEOLKAVMDDDFAAFVEKCKCCKADDDKTCFAEEGKKLVAASQAALGL 585  
DB KEOLKAVMDDDFAAFVEKCKCCKADDDKTCFAEEGKKLVAASQAALGL 585

RESULT 13  
AAE08578  
ID AAE08578 standard; protein; 585 AA.  
XX AAE08578;  
AC AAE08578;  
XX 19-NOV-2001 (first entry)  
DT 19-NOV-2001 (first entry)  
XX Human serum albumin (HSA).  
DE Human; albumin; cancer; cell proliferation; drug screening; biopsy.  
XX Homo sapiens.  
OS Homo sapiens.  
XX US6274305-B1.  
PN US6274305-B1.  
XX 14-AUG-2001.  
PD 14-AUG-2001.  
XX 19-DEC-1996; 96US-00769746.  
PF 19-DEC-1996; 96US-00769746.  
XX 19-DEC-1996; 96US-00769746.  
PR 19-DEC-1996; 96US-00769746.  
XX (TUFT ) UNIV TUFTS.  
PA (TUFT ) UNIV TUFTS.  
XX Sonnenschein C. Soto AM;  
PI Sonnenschein C. Soto AM;  
XX WPI: 2001-540371/60.  
DR N-PSDB; AAD11488.  
XX Measuring human cell proliferation, useful in drug screening to determine  
PT the potential for inhibiting cancer cell proliferation and for evaluating  
PT biopsied tumors, comprises employing albumin-derived peptide.  
XX Claim 1; Fig 1; 20pp; English.  
PS Claim 1; Fig 1; 20pp; English.  
XX The invention related to a method for testing cancer cells. The method is  
CC useful for measuring human cancer cell proliferation, particularly for  
CC determining the potential for inhibiting cancer cells proliferation using  
CC albumin-derived peptides. The invention is also useful for drug screening  
CC assays, as well as for evaluating biopsied tumours. The present sequence  
CC is human serum albumin (HSA) related to the invention  
XX Sequence 585 AA;  
SQ Sequence 585 AA;

Query Match 100.0%; Score 3103; DB 4; Length 585;  
Best Local Similarity 100.0%; Pred. No. 1.9e-254;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFDKLGSENFKALVIAFAQYLOQCPEDHVKLVNEVTEFAKTCVADESAE 60  
DB 1 DAHKEVAHRFDKLGSENFKALVIAFAQYLOQCPEDHVKLVNEVTEFAKTCVADESAE 60  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
QY 121 DWKCTAFHNEETFLKKLYEYARRHPYFYAPPELLUFFAKRYKAAFTCCQADKACLLP 180  
DB 121 DWKCTAFHNEETFLKKLYEYARRHPYFYAPPELLUFFAKRYKAAFTCCQADKACLLP 180  
QY 181 KLDELDEGKASSAKORLKASLOKQGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 240  
DB 181 KLDELDEGKASSAKORLKASLOKQGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 240

QY 241 VHTECCHGDLLECADDRAADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300  
DB 241 VHTECCHGDLLECADDRAADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300  
QY 301 DLPSTLAADVESKDVCKNYAEAKDVFGLMFLYEAARRHPDYSVVLRLAKTYETTLEK 360  
DB 301 DLPSTLAADVESKDVCKNYAEAKDVFGLMFLYEAARRHPDYSVVLRLAKTYETTLEK 360  
QY 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCSELFQGLGEYKFNQALLVRYTKKVPQVST 420  
DB 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCSELFQGLGEYKFNQALLVRYTKKVPQVST 420  
QY 421 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEDVLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 421 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEDVLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
QY 481 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVKKPKAT 540  
DB 481 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVKKPKAT 540  
QY 541 KEOLKAVMDDDFAAFVEKCKCCKADDDKTCFAEEGKKLVAASQAALGL 585  
DB 541 KEOLKAVMDDDFAAFVEKCKCCKADDDKTCFAEEGKKLVAASQAALGL 585

RESULT 14  
AAU75220  
ID AAU75220 standard; protein; 585 AA.  
XX AAU75220;  
AC AAU75220;  
XX 21-MAY-2002 (first entry)  
DT 21-MAY-2002 (first entry)  
XX Mature form of human serum albumin (HSA or HA).  
DE Mature form of human serum albumin (HSA or HA).  
XX Albumin fusion protein; therapeutic protein; immune disorder;  
KW autoimmune disorder; blood-related disorder; hyperproliferative disorder;  
KW renal disorder; cardiovascular disorder; respiratory disorder;  
KW neurological disorder; endocrine disorder; reproductive system disorder;  
KW gastrointestinal disorder; infectious disease; wound healing;  
KW human serum albumin; HSA; HA.  
XX Homo sapiens.  
OS Homo sapiens.  
XX Key Location/Qualifiers  
FH Protein 1..585  
FT /label= Mature\_HSA  
XX WO200179271-A1.  
XX 25-OCT-2001.  
XX 12-APR-2001; 2001WO-US012009.  
XX 12-APR-2000; 2000US-0229358P.  
PR 25-APR-2000; 2000US-0199384P.  
XX 21-DEC-2000; 2000US-0256931P.  
XX (PRIN-) PRINCIPIA PHARM CORP.  
PA (DELZ ) DELTA BIOTECHNOLOGY LTD.  
XX Ballance DJ, Sleep D, Turner AJ, Sadeghi H, Prior CP;  
PI WPI: 2002-179329/23.  
XX N-PSDB; ABK13862.  
XX New albumin fusion proteins with extended shelf life, useful for treating  
PT leukemia, warts, hepatitis, multiple sclerosis and AIDS, comprises  
PT therapeutic protein fused to albumin.  
XX Claim 1; Fig 15; 338pp; English.  
PS The present invention relates to albumin fusion proteins comprising a  
CC

CC therapeutic protein and albumin (e.g. human serum albumin (HSA or HA)).  
CC The albumin fusion proteins are useful for treating, preventing, or  
CC ameliorating various disorders. Such disorders include immune disorders,  
CC autoimmune disorders, blood-related disorders, hyperproliferative  
CC disorders, renal disorders, cardiovascular disorders, respiratory  
CC disorders, neurological disorders, endocrine disorders, reproductive  
CC system disorders. Gastrointestinal disorders, infectious disease, and  
CC wound healing. Therapeutic proteins can be stabilised to extend shelf  
CC life and/or retain the protein's activity for extended periods of time in  
CC solution, in vivo or in vitro by genetically or chemically fusing the  
CC protein to albumin or its fragment or variant. In addition the use of  
CC albumin fusion proteins reduces the need to formulate protein solutions  
CC with large excesses of carrier proteins to prevent loss of therapeutic  
CC protein due to factors such as binding to the container. The extension of  
CC shelf life was tested by measuring biological activity (Nb2 cell  
CC proliferation) of human albumin-human growth hormone (HA-hGH) fusion  
CC protein remaining after incubation in cell culture media for up to 3  
CC weeks. At week 3 there was still approximately 95% cell proliferation  
CC compared to no activity of unfused hGH. The present sequence represents  
CC the mature form of HSA which can be used to produce the albumin fusion  
CC proteins of the invention  
XX

SQ Sequence 585 AA;

Query Match 100.0%; Score 3103; DB 5; Length 585;  
Best Local Similarity 100.0%; Pred. No. 1.9e-254;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGEENFKALVLIAPAYQLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
DB 1 DAHSEVAHRFKDLGEENFKALVLIAPAYQLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPLRLVRPEV 120  
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPLRLVRPEV 120

QY 121 DVMCTAFHDNSETFLKYLIEIARRHPYFYAPPELLFFAKRYKAAVTECCQAADKAACLLP 180  
DB 121 DVMCTAFHDNSETFLKYLIEIARRHPYFYAPPELLFFAKRYKAAVTECCQAADKAACLLP 180

QY 181 KLDELDEGKASSAKQRLKCAASLOKFGGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240  
DB 181 KLDELDEGKASSAKQRLKCAASLOKFGGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240

QY 241 VHTECHGDLLECAADRADLAKYICENQDSISSKLKECEKPLEKSHCIAEVNDENMPA 300  
DB 241 VHTECHGDLLECAADRADLAKYICENQDSISSKLKECEKPLEKSHCIAEVNDENMPA 300

QY 301 DLPSLAADFVSKDKVCKNYAEAKDVFGLGMFLYEYARRHPDYSVLLLLRLAKTYETTLK 360  
DB 301 DLPSLAADFVSKDKVCKNYAEAKDVFGLGMFLYEYARRHPDYSVLLLLRLAKTYETTLK 360

QY 361 CAADPHCYAKVDFEKPPLVEEPQNLIKQNCLEPQGEYKFQNALVRYTKVPQVST 420  
DB 361 CAADPHCYAKVDFEKPPLVEEPQNLIKQNCLEPQGEYKFQNALVRYTKVPQVST 420

QY 421 PTLVEVSNLGVSKGCKCHPEAKMPCAEADYLSVLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 421 PTLVEVSNLGVSKGCKCHPEAKMPCAEADYLSVLNQLCVLHEKTPVSDRVTKCCTES 480

QY 481 LVNRRPCFSALEVDYTPVPEFNAETFTFHADICTLSEKERQIKQTALVELVGHKPKAT 540  
DB 481 LVNRRPCFSALEVDYTPVPEFNAETFTFHADICTLSEKERQIKQTALVELVGHKPKAT 540

QY 541 KEQLKAVMDDFAAFVEKCKADDKTCFAEBGKKLVAAASQAALGL 585  
DB 541 KEQLKAVMDDFAAFVEKCKADDKTCFAEBGKKLVAAASQAALGL 585

RESULT 15  
ABJ00986  
ID ABJ00986 standard; protein; 585 AA.  
XX

AC ABJ00986;  
XX  
DT 05-SEP-2002 (first entry)  
XX  
DE B lymphocyte stimulator protein binding peptide related protein.  
XX  
DE B lymphocyte stimulator protein binding protein; Blys; immune disease;  
KW allergy; proliferative disease; infectious disease; arteriosclerosis;  
KW inflammatory disorder; hypergammaglobulinaemia; blood clotting;  
KW ischaemia; graft-versus-host disease; neurodegenerative disease;  
KW immunosuppressive; nephrotropic; antirheumatic; antiarthritic;  
KW neuroproliferative; cytostatic; immunostimulant; antitumour; anti-HIV;  
KW antiasthmatic; antiallergic; thymomimetic; antianemic; haemostatic;  
KW dermatological; antiinflammatory; cardiant; ophthalmological; uropathic;  
KW antidiabetic; antithyroid; antidepressant; hepatotropic.  
XX  
OS Homo sapiens.  
XX  
OS WO200216411-A2.  
PN  
XX 28-FEB-2002.  
PD  
XX  
XX 17-AUG-2001; 2001WO-0258550.  
PF  
XX 18-AUG-2000; 2000US-0226700P.  
PR  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA  
XX Beltzer JP, Potter DM, Fleming TL, Rosen CA;  
PI  
XX WPI; 2002-499775/53.  
DR  
XX  
XX The treatment of various diseases e.g. rheumatoid arthritis, comprises  
PT administering B lymphocyte stimulator binding polypeptide.  
PT  
XX  
PS Disclosure; Page 379-382; 387pp; English.  
XX  
CC The present invention relates to the treatment, prevention or  
CC amelioration of a disease or disorder associated with: aberrant B  
CC lymphocyte stimulator (Blys), Blys receptor expression or activity; cells  
CC of haematopoietic origin; or proliferative disease; and reducing,  
CC inhibiting or stimulating immunoglobulin production, B cell proliferation  
CC and graft rejection involving administration of Blys binding polypeptide.  
CC The Blys binding polypeptides are used in the treatment, prevention or  
CC amelioration of diseases such as immune system diseases, proliferative  
CC diseases, diseases of cells of hematopoietic origin, graft rejection,  
CC allergies, infectious diseases, arteriosclerosis, inflammatory disorders,  
CC hypergammaglobulinaemia, blood clotting disorders, ischaemia, and  
CC neurodegenerative diseases. The present sequence is a protein described  
CC in the invention  
XX  
SQ Sequence 585 AA;

Query Match 100.0%; Score 3103; DB 5; Length 585;  
Best Local Similarity 100.0%; Pred. No. 1.9e-254;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGEENFKALVLIAPAYQLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
DB 1 DAHSEVAHRFKDLGEENFKALVLIAPAYQLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPLRLVRPEV 120  
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPLRLVRPEV 120

QY 121 DVMCTAFHDNSETFLKYLIEIARRHPYFYAPPELLFFAKRYKAAVTECCQAADKAACLLP 180  
DB 121 DVMCTAFHDNSETFLKYLIEIARRHPYFYAPPELLFFAKRYKAAVTECCQAADKAACLLP 180

QY 181 KLDELDEGKASSAKQRLKCAASLOKFGGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240  
DB 181 KLDELDEGKASSAKQRLKCAASLOKFGGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240



XX 28-FEB-2002.  
XX  
XX  
XX 17-AUG-2001; 2001WO-US025891.  
XX  
XX 18-AUG-2000; 2000US-0226489P.  
XX  
XX (DYAX-) DVAX CORP.  
XX  
XX Beltzer JP, Potter MD, Fleming TJ, Ladner RC;  
XX WPI; 2002-351647/38.  
XX  
XX New B-lymphocyte stimulator binding polypeptide useful in detecting or  
XX isolating BlyS or BlyS-like polypeptide comprises a specified amino acid  
XX sequence.  
XX  
XX Disclosure; Page 261-264; 269pp; English.  
XX  
XX The invention relates to a B lymphocyte stimulator (BlyS) binding  
XX polypeptide. BlyS binding peptides bind BlyS or BlyS-like proteins  
XX reversibly or irreversibly. The binding peptides are used in detection,  
XX isolation and/or purification of BlyS in a solution such as water or a  
XX buffer solution, as well as any fluid and/or cell obtained from an  
XX individual biological fluid, body tissue, body cell, cell line, tissue  
XX culture or other source containing BlyS or BlyS-like polypeptides. The  
XX biological fluids include sera, plasma, lymph, blood, blood fraction,  
XX urine, synovial fluid, spinal fluid, saliva and mucus. Sequences  
XX ABG33576, ABG33577 and ABG33847 represent human B lymphocyte stimulator  
XX proteins  
XX  
XX Sequence 585 AA;  
XX  
Query Match 100.0%; Score 3103; DB 5; Length 585;  
Best Local Similarity 100.0%; Pred. No. 1.9e-254;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
QY 1 DAHSEVAHRFKDLGEENFKALVLIATAFYAQYLOQCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
DB 1 DAHSEVAHRFKDLGEENFKALVLIATAFYAQYLOQCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPMLPRLVRPEV 120  
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPMLPRLVRPEV 120  
QY 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLLP 180  
DB 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLLP 180  
QY 181 KLDELRDGKASSAKQRLKCSAQKQFGERAFKAWAVARLSORFPKAEAEVSKLVTDLTK 240  
DB 181 KLDELRDGKASSAKQRLKCSAQKQFGERAFKAWAVARLSORFPKAEAEVSKLVTDLTK 240  
QY 241 VHTCCHGDLLECADRADLAKYICENQDSISSKLKECCERPLEKSHCIAEVENDEMPA 300  
DB 241 VHTCCHGDLLECADRADLAKYICENQDSISSKLKECCERPLEKSHCIAEVENDEMPA 300  
QY 301 DLPSLAADFVSKDKVCKNYAEAKDVLGMFLYFYARRHPDYSVLLLRKAKTYETTLK 360  
DB 301 DLPSLAADFVSKDKVCKNYAEAKDVLGMFLYFYARRHPDYSVLLLRKAKTYETTLK 360  
QY 361 CAAADPHCYAKVDFEPLVEEPQNLKQNCLEFQGEYKFNALLVRYTKVPQVST 420  
DB 361 CAAADPHCYAKVDFEPLVEEPQNLKQNCLEFQGEYKFNALLVRYTKVPQVST 420  
QY 421 PTLVEVSRNLGVSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 421 PTLVEVSRNLGVSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
QY 481 LVNRPCTFSALVDETYVPKFNATFTPHADICTLSEKERQIKQTALVELVKKPKAT 540  
DB 481 LVNRPCTFSALVDETYVPKFNATFTPHADICTLSEKERQIKQTALVELVKKPKAT 540

QY 541 KEQLKAVMDDFAAFVEKCKCKADDDKTCFABEGKKLVAASQAALGL 585  
DB 541 KEQLKAVMDDFAAFVEKCKCKADDDKTCFABEGKKLVAASQAALGL 585  
RESULT 18  
ABG71291  
ID ABG71291 standard; protein; 585 AA.  
XX  
XX ABG71291;  
AC  
XX 08-JAN-2003 (first entry)  
DT  
XX Glycosylated protein determination associated protein.  
DE  
XX Standard substance; accuracy control substance; glycosylated protein;  
KW glycosylated albumin; fructosamine; diabetes; antidiabetic.  
XX  
XX Unidentified.  
OS  
XX JP2002243731-A.  
PN  
XX 28-AUG-2002.  
PD  
XX 21-FEB-2001; 2001JP-00045085.  
PF  
XX 21-FEB-2001; 2001JP-00045085.  
PR  
XX (KOKU-) KOKUSAI SHYAKU KK.  
XX (YOSH-) YOSHITOMI PHARM IND KK.  
PA  
XX WPI; 2002-744850/81.  
DR  
XX  
XX A standard substance for determination of glycosylated protein including  
PT glycosylated albumin and fructosamine, used in diagnosis of diabetes.  
XX  
XX Disclosure; Page 4; 6pp; Japanese.  
XX  
XX The present invention relates to a new standard and accuracy control  
XX substance for determination of glycosylated protein. The invention is  
XX useful for determination of glycosylated protein in the diagnosis of  
XX diabetes. Glycosylated albumin and fructosamine provide favourable  
XX dilution linearity. The present amino acid sequence represents the  
XX glycosylated protein determination associated protein as described in the  
XX invention  
XX  
XX Sequence 585 AA;  
XX  
Query Match 100.0%; Score 3103; DB 5; Length 585;  
Best Local Similarity 100.0%; Pred. No. 1.9e-254;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DAHSEVAHRFKDLGEENFKALVLIATAFYAQYLOQCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
DB 1 DAHSEVAHRFKDLGEENFKALVLIATAFYAQYLOQCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPMLPRLVRPEV 120  
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPMLPRLVRPEV 120  
QY 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLLP 180  
DB 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLLP 180  
QY 181 KLDELRDGKASSAKQRLKCSAQKQFGERAFKAWAVARLSORFPKAEAEVSKLVTDLTK 240  
DB 181 KLDELRDGKASSAKQRLKCSAQKQFGERAFKAWAVARLSORFPKAEAEVSKLVTDLTK 240  
QY 241 VHTCCHGDLLECADRADLAKYICENQDSISSKLKECCERPLEKSHCIAEVENDEMPA 300  
DB 241 VHTCCHGDLLECADRADLAKYICENQDSISSKLKECCERPLEKSHCIAEVENDEMPA 300  
QY 301 DLPSLAADFVSKDKVCKNYAEAKDVLGMFLYFYARRHPDYSVLLLRKAKTYETTLK 360

Db 301 DLPSSLAADFVESKDVCKNYAEAKDVFGLGMFLYEYARRHPDYSVLLRLAKTYETTTLEK 360  
QY 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCFLFEQLGEYKFNALLVRYTKKVPQVST 420  
Db 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCFLFEQLGEYKFNALLVRYTKKVPQVST 420  
QY 421 PTLVEVSRNLGKVGSKCCCKHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
Db 421 PTLVEVSRNLGKVGSKCCCKHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTPHADICTLSEKEROIKKOTALVELVGHKPKAT 540  
Db 481 LVNRRPCFSALEVDETYVPKEFNAETFTPHADICTLSEKEROIKKOTALVELVGHKPKAT 540  
QY 541 KEQLKAVMDDDFAAFVEKCKCKADDKETCFAEEGKLVAAASQAALGL 585  
Db 541 KEQLKAVMDDDFAAFVEKCKCKADDKETCFAEEGKLVAAASQAALGL 585

RESULT 19  
ID ABR55695  
XX ABR55695 standard; protein; 585 AA.  
AC ABR55695;  
XX  
XX  
DT 18-AUG-2003 (first entry)  
XX  
DE Human albumin (HA) mature polypeptide.  
XX  
KW Albumin; HA; cytostatic; antibacterial; virucide; fungicide; anti-HIV;  
KW antiasthmatic; osteopathic; antiarthritic; antiinflammatory; nootropic;  
KW neuroprotective; anti-thyroid; anti-ulcer; hepatotropic; vulnerary;  
XX  
XX Homo sapiens.  
XX  
PN WO2003030821-A2.  
XX  
PD 17-APR-2003.  
XX  
XX 04-OCT-2002; 2002WO-US031794.  
PF  
XX  
PR 05-OCT-2001; 2001US-0327281P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Haseeltine WA;  
XX  
XX WPI; 2003-468174/44.  
DR N-PSDB; ACC78574.  
XX  
XX New albumin fusion protein comprising a therapeutic protein:X, and  
PT albumin, its variant or fragment, useful for treating a cancer, AIDS  
PT asthma, leukemia, sepsis, endometriosis, osteoporosis, atherosclerosis,  
PT autism, or emphysema.  
XX  
PS Claim 1; Fig 15A-D; 455pp; English.  
XX  
XX The invention relates to an albumin fusion protein comprising a  
CC therapeutic protein:X, and albumin, its variant or fragment. The albumin  
CC fusion protein has the formula R1-L-R2; R2-L-R1; or R1-L-R2-L-R1 where R1  
CC is therapeutic protein:X or fragment, L is a peptide linker and R2 is  
CC albumin. The albumin fusion protein is useful for treating a disease or  
CC disorder that is modulated by therapeutic protein:X (claimed), such as  
CC cancer; infections (bacterial, viral, fungal, parasitic); or immune  
CC (AIDS, asthma); hematopoietic (leukemia, sepsis); reproductive (cystic  
CC fibrosis, endometriosis); musculoskeletal (osteoporosis, osteoarthritis);  
CC cardiovascular (congestive heart failure, atherosclerosis); neural/  
CC sensory (ataxia, attention deficit disorders, autism); respiratory  
CC (emphysema, bronchitis); endocrine (goiter, glomerulonephritis);  
CC digestive (ulcer, cirrhosis); or connective/epithelial (lupus, keloids)  
CC disorders. The present sequence represents a human albumin (HA) mature

CC polypeptide  
XX SQ Sequence 585 AA;  
Query Match 100.0%; Score 3103; DB 6; Length 585;  
Best Local Similarity 100.0%; Pred. No. 1.9e-254; Indels 0; Gaps 0;  
Matches 585; Conservative 0; Mismatches 0;  
QY 1 DAHSEVAHRFKDLGEENFKALVLIATAFYLOQCPEFDHVKLVNEVTEFAKTCVADESAAE 60  
Db 1 DAHSEVAHRFKDLGEENFKALVLIATAFYLOQCPEFDHVKLVNEVTEFAKTCVADESAAE 60  
QY 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQBERNECFLOHKDDNPRLVRPEV 120  
Db 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQBERNECFLOHKDDNPRLVRPEV 120  
QY 121 DWCTAFHDNEETFLKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 180  
Db 121 DWCTAFHDNEETFLKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 180  
QY 181 KLDELDEGKASSAKQRLKCSLQKGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 240  
Db 181 KLDELDEGKASSAKQRLKCSLQKGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 240  
QY 241 VHTTECHGDLLECCADDRADLAKYICENQDSISSKLUKCECKPLLEKSHGICAEVENDMPA 300  
Db 241 VHTTECHGDLLECCADDRADLAKYICENQDSISSKLUKCECKPLLEKSHGICAEVENDMPA 300  
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYEYARRHPDYSVLLRLAKTYETTTLEK 360  
Db 301 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYEYARRHPDYSVLLRLAKTYETTTLEK 360  
QY 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCFLFEQLGEYKFNALLVRYTKKVPQVST 420  
Db 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCFLFEQLGEYKFNALLVRYTKKVPQVST 420  
QY 421 PTLVEVSRNLGKVGSKCCCKHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
Db 421 PTLVEVSRNLGKVGSKCCCKHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTPHADICTLSEKEROIKKOTALVELVGHKPKAT 540  
Db 481 LVNRRPCFSALEVDETYVPKEFNAETFTPHADICTLSEKEROIKKOTALVELVGHKPKAT 540  
QY 541 KEQLKAVMDDDFAAFVEKCKCKADDKETCFAEEGKLVAAASQAALGL 585  
Db 541 KEQLKAVMDDDFAAFVEKCKCKADDKETCFAEEGKLVAAASQAALGL 585  
RESULT 20  
ID ABR42606  
XX ABR42606 standard; protein; 585 AA.  
XX  
AC ABR42606;  
XX  
DT 26-AUG-2003 (first entry)  
XX  
DE Human serum albumin.  
XX  
KW Human; serum albumin; abrogen; angiogenesis; inhibitor;  
KW urokinase plasminogen activator; tumour; metastasis; cytostatic;  
KW gene therapy.  
OS Homo sapiens.  
XX  
PN WO2003042354-A2.  
PD 22-MAY-2003.  
XX  
PF 04-SEP-2002; 2002WO-US027885.  
XX  
PR 04-SEP-2001; 2001US-0316300P.  
XX



PA (AVET ) AVENTIS PHARM INC.  
XX  
PI Nesbit M, Fong TC, Brockstedt D;  
XX  
DR WPI; 2003-449566/42.  
XX  
PT New abrogen polypeptide, useful for treating an angiogenesis related  
XX diseases e.g. tumor metastasis.  
XX  
PS Disclosure; Page 26; 95pp; English.  
XX  
CC The present sequence is the protein sequence of human serum albumin  
CC (HSA) . The invention provides novel abrogen polypeptides (see ABR42599-  
CC 602) that include a kringle domain, e.g. from urokinase plasminogen  
CC activator kringle domain. Abrogens are potent inhibitors of endothelial  
CC proliferation and angiogenesis, and have been shown to be capable of  
CC inhibiting or reducing cell proliferation induced by both basic  
CC fibroblast growth factor and vascular endothelial growth factor in a  
CC specific endothelial cell proliferation assay. The abrogen may be coupled  
CC to an N-terminal interleukin-2 signal peptide and to a C-terminal  
CC stabilising molecule such as HSA (see ABR42608-13). Abrogen polypeptides  
CC and polynucleotides are used in claimed methods of treating an  
CC angiogenesis-related disease or disorder, e.g. tumour metastasis  
XX  
SQ Sequence 585 AA;  
  
Query Match 100.0%; Score 3103; DB 7; Length 585;  
Best Local Similarity 100.0%; Pred. No. 1.9e-254;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 DAHKEVAHRFKDLGEENFKALVLIAPAYQLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
DB 1 DAHKEVAHRFKDLGEENFKALVLIAPAYQLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
  
QY 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPMLPRLVRPEV 120  
DB 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPMLPRLVRPEV 120  
  
QY 121 DVMCTAFHDNBEFTFLKKYLIEIARHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
DB 121 DVMCTAFHDNBEFTFLKKYLIEIARHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
  
QY 181 KLDELDRGKASSAKQRLKCSLQKFGERAPKAWAVARLSORFPKAEFAEVSCLVTDLTG 240  
DB 181 KLDELDRGKASSAKQRLKCSLQKFGERAPKAWAVARLSORFPKAEFAEVSCLVTDLTG 240  
  
QY 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKLKECEKPLLEKSHCIAEVENDENMPA 300  
DB 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKLKECEKPLLEKSHCIAEVENDENMPA 300  
  
QY 301 DLPSLAADPVESKDYCKNYAEAKDVFGLGMYEYARRHPDYSVLLLRLLAKTYETTTLEKC 360  
DB 301 DLPSLAADPVESKDYCKNYAEAKDVFGLGMYEYARRHPDYSVLLLRLLAKTYETTTLEKC 360  
  
QY 361 CAADPHECYAKVDFEFPKPLVEEPONLIKQNCLEFEQLGEYKFQNALLVRYTKVPQVST 420  
DB 361 CAADPHECYAKVDFEFPKPLVEEPONLIKQNCLEFEQLGEYKFQNALLVRYTKVPQVST 420  
  
QY 421 PTLVEVSRNLGKVGSKCKCKHPEAKMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 421 PTLVEVSRNLGKVGSKCKCKHPEAKMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
  
QY 481 LVNRRPCFSALVEDTTPYVPEKFNAEFTTFHADICTLSEKERQIKKQTALVELVKKPKAT 540  
DB 481 LVNRRPCFSALVEDTTPYVPEKFNAEFTTFHADICTLSEKERQIKKQTALVELVKKPKAT 540  
  
QY 541 KEQLKAVMDDFAAFVEKCKKADDTCTFAEGGKLVAAASQAALGL 585  
DB 541 KEQLKAVMDDFAAFVEKCKKADDTCTFAEGGKLVAAASQAALGL 585



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2005, 08:42:51 ; Search time 25 Seconds  
(without alignments)  
1746.788 Million cell updates/sec

Title: US-10-816-042-18  
Perfect score: 3103  
Sequence: 1 DHAKEVAHFKDLGEENPK.....TCFAEKGKLVAAQAALGL 585

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/prodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/prodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/prodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/prodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/prodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3103	100.0	585	1	US-08-153-799-14
2	3103	100.0	585	2	US-08-702-572-2
3	3103	100.0	585	3	US-08-769-746-2
4	3103	100.0	609	4	US-09-976-594-977
5	3103	100.0	609	4	US-09-919-039-370
6	3103	100.0	610	2	US-08-797-689-2
7	3103	100.0	610	4	US-09-984-186-2
8	3103	100.0	622	4	US-09-949-016-11170
9	3103	100.0	783	1	US-08-256-938-2
10	3103	100.0	787	1	US-08-256-938-4
11	3103	100.0	787	2	US-08-797-689-16
12	3103	100.0	787	4	US-09-984-186-2
13	3099	99.9	609	1	US-08-222-619-3
14	3099	99.9	609	1	US-08-433-037-4
15	3099	99.9	609	4	US-08-897-956A-2
16	3099	99.9	609	5	PCT-US95-04075-3
17	3099	99.7	778	4	US-08-897-956A-3
18	3093	99.7	585	1	US-08-448-196A-3
19	3093	99.7	585	2	US-08-984-176-1
20	2458.5	79.2	583	1	US-08-448-196A-5
21	2450.5	79.0	583	1	US-08-448-196A-4
22	2432.5	78.4	583	1	US-08-448-196A-6
23	2426	78.2	584	1	US-08-448-196A-7
24	2389	77.0	582	1	US-08-134-638-1
25	1249.5	40.3	609	1	US-08-222-619-4
26	1249.5	40.3	609	4	US-09-976-594-456
27	1249.5	40.3	609	5	PCT-US95-04075-4

28	1249.5	40.3	612	4	US-09-949-016-11201	Sequence 11201, A
29	1206.5	38.9	590	2	US-08-377-309-2	Sequence 2, Appli
30	1206.5	38.9	590	3	US-09-186-723-2	Sequence 2, Appli
31	1206.5	38.9	590	3	US-08-505-012-5	Sequence 5, Appli
32	1206.5	38.9	590	4	US-09-186-949A-3	Sequence 3, Appli
33	1206.5	38.9	590	4	US-08-758-757-2	Sequence 2, Appli
34	1206.5	38.9	590	4	US-09-187-978-2	Sequence 2, Appli
35	1206.5	38.9	590	4	US-10-115-701A-2	Sequence 2, Appli
36	1206.5	38.9	590	4	US-09-940-308A-2	Sequence 2, Appli
37	1206.5	38.9	590	4	US-09-940-308A-2	Sequence 2, Appli
38	1206.5	38.9	590	5	PCT-US96-00996-5	Sequence 5, Appli
39	1206.5	38.9	609	4	US-09-186-949A-2	Sequence 5, Appli
40	1164.5	37.5	579	1	US-08-448-196A-8	Sequence 8, Appli
41	1055	34.0	599	1	US-08-222-619-2	Sequence 2, Appli
42	1055	34.0	599	3	US-08-221-767-24	Sequence 24, Appli
43	1055	34.0	599	5	PCT-US95-04075-2	Sequence 2, Appli
44	926	29.8	393	2	US-08-377-309-7	Sequence 7, Appli
45	926	29.8	393	3	US-09-186-723-7	Sequence 7, Appli
46	926	29.8	393	3	US-08-505-012-10	Sequence 10, Appli
47	926	29.8	393	4	US-09-186-949A-8	Sequence 8, Appli
48	926	29.8	393	4	US-08-758-757-7	Sequence 7, Appli
49	926	29.8	393	4	US-09-187-978-7	Sequence 7, Appli
50	926	29.8	393	4	US-10-115-701A-7	Sequence 7, Appli
51	926	29.8	393	4	US-09-940-308A-7	Sequence 7, Appli
52	926	29.8	393	4	US-09-940-308A-7	Sequence 7, Appli
53	926	29.8	393	5	PCT-US96-00996-10	Sequence 10, Appli
54	777	25.0	324	3	US-08-505-012-12	Sequence 12, Appli
55	777	25.0	324	5	PCT-US96-00996-12	Sequence 12, Appli
56	777	25.0	325	2	US-08-377-309-8	Sequence 8, Appli
57	777	25.0	325	3	US-09-186-723-8	Sequence 8, Appli
58	777	25.0	325	3	US-08-505-012-11	Sequence 11, Appli
59	777	25.0	325	4	US-09-186-949A-9	Sequence 9, Appli
60	777	25.0	325	4	US-08-758-757-8	Sequence 8, Appli
61	777	25.0	325	4	US-09-187-978-8	Sequence 8, Appli
62	777	25.0	325	4	US-10-115-701A-8	Sequence 8, Appli
63	777	25.0	325	4	US-09-940-308A-8	Sequence 8, Appli
64	777	25.0	325	4	US-09-940-308A-8	Sequence 8, Appli
65	777	25.0	325	5	PCT-US96-00996-11	Sequence 11, Appli
66	747.5	24.1	590	1	US-08-448-196A-9	Sequence 9, Appli
67	684.5	22.1	389	2	US-08-377-309-6	Sequence 6, Appli
68	684.5	22.1	389	3	US-09-186-723-6	Sequence 6, Appli
69	684.5	22.1	389	3	US-08-505-012-9	Sequence 9, Appli
70	684.5	22.1	389	4	US-09-186-949A-7	Sequence 7, Appli
71	684.5	22.1	389	4	US-08-758-757-6	Sequence 6, Appli
72	684.5	22.1	389	4	US-09-187-978-6	Sequence 6, Appli
73	684.5	22.1	389	4	US-10-115-701A-6	Sequence 6, Appli
74	684.5	22.1	389	4	US-09-940-308A-6	Sequence 6, Appli
75	684.5	22.1	389	4	US-09-940-308A-6	Sequence 6, Appli
76	684.5	22.1	389	5	PCT-US96-00996-9	Sequence 9, Appli
77	523	16.9	201	2	US-08-377-309-5	Sequence 5, Appli
78	523	16.9	201	3	US-09-186-723-5	Sequence 5, Appli
79	523	16.9	201	3	US-08-505-012-8	Sequence 8, Appli
80	523	16.9	201	4	US-09-186-949A-6	Sequence 6, Appli
81	523	16.9	201	4	US-08-758-757-5	Sequence 5, Appli
82	523	16.9	201	4	US-09-187-978-5	Sequence 5, Appli
83	523	16.9	201	4	US-10-115-701A-5	Sequence 5, Appli
84	523	16.9	201	4	US-09-940-308A-5	Sequence 5, Appli
85	523	16.9	201	4	US-09-940-308A-5	Sequence 5, Appli
86	523	16.9	201	5	PCT-US96-00996-8	Sequence 8, Appli
87	521	16.8	117	1	US-08-448-196A-1	Sequence 1, Appli
88	516	16.6	109	1	US-08-448-196A-1	Sequence 1, Appli
89	404	13.0	192	2	US-08-377-309-4	Sequence 4, Appli
90	404	13.0	192	3	US-09-186-723-4	Sequence 4, Appli
91	404	13.0	192	3	US-08-505-012-7	Sequence 7, Appli
92	404	13.0	192	4	US-09-186-949A-5	Sequence 5, Appli
93	404	13.0	192	4	US-08-758-757-4	Sequence 4, Appli
94	404	13.0	192	4	US-09-187-978-4	Sequence 4, Appli
95	404	13.0	192	4	US-10-115-701A-4	Sequence 4, Appli
96	404	13.0	192	4	US-09-940-308A-4	Sequence 4, Appli
97	404	13.0	192	4	US-09-940-308A-4	Sequence 4, Appli
98	404	13.0	192	5	PCT-US96-00996-7	Sequence 7, Appli
99	388	12.5	474	4	US-09-949-016-11630	Sequence 11630, A
100	386	12.4	458	4	US-08-618-485B-1	Sequence 1, Appli



ATTORNEY/AGENT INFORMATION:  
NAME: Naomi Biswas  
REGISTRATION NUMBER: 38,384  
REFERENCE/DOCKET NUMBER: CE0114 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610/878/4294  
TELEFAX: 610/878/4221  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 585 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-702-572-2

Query Match 100.0%; Score 3103; DB 2; Length 585;  
Best Local Similarity 100.0%; Pred. No. 6.5e-287;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFKDLGEENFKALVLIATAFYLOQCPEDHVKLVNEVTEFAKTCVADESAE 60  
DB 1 DAHKEVAHRFKDLGEENFKALVLIATAFYLOQCPEDHVKLVNEVTEFAKTCVADESAE 60

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKKDDNPNLRLVRPEV 120  
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKKDDNPNLRLVRPEV 120

QY 121 DVMCTAFHNDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLIP 180  
DB 121 DVMCTAFHNDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLIP 180

QY 181 KLDELDEGKASAKQRLKCSLQKFGGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 240  
DB 181 KLDELDEGKASAKQRLKCSLQKFGGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 240

QY 241 VHTCCHGDLLECADDRAADLAKYICENQDSISSKLKECCCKPLEKSHCHIAEVENDEMPA 300  
DB 241 VHTCCHGDLLECADDRAADLAKYICENQDSISSKLKECCCKPLEKSHCHIAEVENDEMPA 300

QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLRLAKTYETTLK 360  
DB 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLRLAKTYETTLK 360

QY 361 CAADPHCEYAKVDFEFKPLVEEPQNLIKONCELFEQLGEYKFNALLVRYTKVPQVST 420  
DB 361 CAADPHCEYAKVDFEFKPLVEEPQNLIKONCELFEQLGEYKFNALLVRYTKVPQVST 420

QY 421 PTLVEVSRNLGVKSGCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 421 PTLVEVSRNLGVKSGCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480

QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKGHPKAT 540  
DB 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKGHPKAT 540

QY 541 KEQLKAVMDDDFAAFVEKCKCKADDKETCFABEKGKLVAAQAALGL 585  
DB 541 KEQLKAVMDDDFAAFVEKCKCKADDKETCFABEKGKLVAAQAALGL 585

RESULT 3  
US-08-769-746-2  
Sequence 2, Application US/08769746  
Patent No. 6274305  
GENERAL INFORMATION:  
APPLICANT: Sonnenschein, Carlos  
APPLICANT: Soto, Ana M.  
TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSES: Medien & Carroll, LLP  
STREET: 220 Montgomery Street, Suite 2200  
CITY: San Francisco

STATE: California  
COUNTRY: United States of America  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/769,746  
FILING DATE: 19-DEC-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Carroll, Peter G.  
REGISTRATION NUMBER: 32,837  
REFERENCE/DOCKET NUMBER: MBRI-02584  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 705-8410  
TELEFAX: (415) 397-8338  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 585 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-769-746-2

Query Match 100.0%; Score 3103; DB 3; Length 585;  
Best Local Similarity 100.0%; Pred. No. 6.5e-287;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFKDLGEENFKALVLIATAFYLOQCPEDHVKLVNEVTEFAKTCVADESAE 60  
DB 1 DAHKEVAHRFKDLGEENFKALVLIATAFYLOQCPEDHVKLVNEVTEFAKTCVADESAE 60

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKKDDNPNLRLVRPEV 120  
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKKDDNPNLRLVRPEV 120

QY 121 DVMCTAFHNDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLIP 180  
DB 121 DVMCTAFHNDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLIP 180

QY 181 KLDELDEGKASAKQRLKCSLQKFGGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 240  
DB 181 KLDELDEGKASAKQRLKCSLQKFGGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 240

QY 241 VHTCCHGDLLECADDRAADLAKYICENQDSISSKLKECCCKPLEKSHCHIAEVENDEMPA 300  
DB 241 VHTCCHGDLLECADDRAADLAKYICENQDSISSKLKECCCKPLEKSHCHIAEVENDEMPA 300

QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLRLAKTYETTLK 360  
DB 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLRLAKTYETTLK 360

QY 361 CAADPHCEYAKVDFEFKPLVEEPQNLIKONCELFEQLGEYKFNALLVRYTKVPQVST 420  
DB 361 CAADPHCEYAKVDFEFKPLVEEPQNLIKONCELFEQLGEYKFNALLVRYTKVPQVST 420

QY 421 PTLVEVSRNLGVKSGCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 421 PTLVEVSRNLGVKSGCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480

QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKGHPKAT 540  
DB 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKGHPKAT 540

QY 541 KEQLKAVMDDDFAAFVEKCKCKADDKETCFABEKGKLVAAQAALGL 585  
DB 541 KEQLKAVMDDDFAAFVEKCKCKADDKETCFABEKGKLVAAQAALGL 585

RESULT 4

```
US-09-976-594-977
; Sequence 977, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 977
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; FEATURE:
; OTHER INFORMATION: Incyte ID No. 6673549 088957CD1
US-09-976-594-977
```

```
Query Match      100.0%; Score 3103; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 6.9e-287;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGEENFKALVLIAPAYIQQCFEDHVKLVNVEVTFBFAKTCVADESAAE 60
Db 25 DAHKSEVAHRFKDLGEENFKALVLIAPAYIQQCFEDHVKLVNVEVTFBFAKTCVADESAAE 84

QY 61 NCDKSLHTLFGDKLCTVATLTRETYGEMADCCAKQEPERNECFLOHKDNDPNLPRLVREPV 120
Db 85 NCDKSLHTLFGDKLCTVATLTRETYGEMADCCAKQEPERNECFLOHKDNDPNLPRLVREPV 144

QY 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELFFAKRYKAAFTCCQAADKAACLLP 180
Db 145 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELFFAKRYKAAFTCCQAADKAACLLP 204

QY 181 KLDELURDGGKASSAKQRLKCSLQKFGGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240
Db 205 KLDELURDGGKASSAKQRLKCSLQKFGGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 264

QY 241 VHTECCGDLLECCADDDRADLAKYICENODSISSKLKECCPKLLEKSHCIAEVENDEMPA 300
Db 265 VHTECCGDLLECCADDDRADLAKYICENODSISSKLKECCPKLLEKSHCIAEVENDEMPA 324

QY 301 DLPSLAADFVBSKDYCKNYAEAKDVFLGMFLYFYARRHPDYSVVLLLRLLAKTYETTTLEKC 360
Db 325 DLPSLAADFVBSKDYCKNYAEAKDVFLGMFLYFYARRHPDYSVVLLLRLLAKTYETTTLEKC 384

QY 361 CAADDPHECYAKVDFEFPKPLVEEPONLIKONCELFEOQLGEYKFQNALLVRYTKKVPQVST 420
Db 385 CAADDPHECYAKVDFEFPKPLVEEPONLIKONCELFEOQLGEYKFQNALLVRYTKKVPQVST 444

QY 421 PTLVEVSRNLGKVGSKCKCHPEAKRMPCAEDYLSVVLNQLCVLHKEKTPVSDRVTKCCTES 480
Db 445 PTLVEVSRNLGKVGSKCKCHPEAKRMPCAEDYLSVVLNQLCVLHKEKTPVSDRVTKCCTES 504

QY 481 LVNRRPCFSALVEDETYVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVVKHKPKAT 540
Db 505 LVNRRPCFSALVEDETYVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVVKHKPKAT 564

QY 541 KEQLKAVMDDFAAFVEKCKCADDKTCFAEBEGKKLVAASQAALGL 585
Db 565 KEQLKAVMDDFAAFVEKCKCADDKTCFAEBEGKKLVAASQAALGL 609
```

```
RESULT 5
US-09-919-039-370
; Sequence 370, Application US/09919039
; Patent No. 6727066
; GENERAL INFORMATION:
```

```
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT APPLICATION NUMBER: US/09/919,039
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/222,113
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 370
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; FEATURE:
; OTHER INFORMATION: Incyte ID No. 6727066 088957CD1
US-09-919-039-370
```

```
Query Match      100.0%; Score 3103; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 6.9e-287;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGEENFKALVLIAPAYIQQCFEDHVKLVNVEVTFBFAKTCVADESAAE 60
Db 25 DAHKSEVAHRFKDLGEENFKALVLIAPAYIQQCFEDHVKLVNVEVTFBFAKTCVADESAAE 84

QY 61 NCDKSLHTLFGDKLCTVATLTRETYGEMADCCAKQEPERNECFLOHKDNDPNLPRLVREPV 120
Db 85 NCDKSLHTLFGDKLCTVATLTRETYGEMADCCAKQEPERNECFLOHKDNDPNLPRLVREPV 144

QY 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELFFAKRYKAAFTCCQAADKAACLLP 180
Db 145 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELFFAKRYKAAFTCCQAADKAACLLP 204

QY 181 KLDELURDGGKASSAKQRLKCSLQKFGGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240
Db 205 KLDELURDGGKASSAKQRLKCSLQKFGGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 264

QY 241 VHTECCGDLLECCADDDRADLAKYICENODSISSKLKECCPKLLEKSHCIAEVENDEMPA 300
Db 265 VHTECCGDLLECCADDDRADLAKYICENODSISSKLKECCPKLLEKSHCIAEVENDEMPA 324

QY 301 DLPSLAADFVBSKDYCKNYAEAKDVFLGMFLYFYARRHPDYSVVLLLRLLAKTYETTTLEKC 360
Db 325 DLPSLAADFVBSKDYCKNYAEAKDVFLGMFLYFYARRHPDYSVVLLLRLLAKTYETTTLEKC 384

QY 361 CAADDPHECYAKVDFEFPKPLVEEPONLIKONCELFEOQLGEYKFQNALLVRYTKKVPQVST 420
Db 385 CAADDPHECYAKVDFEFPKPLVEEPONLIKONCELFEOQLGEYKFQNALLVRYTKKVPQVST 444

QY 421 PTLVEVSRNLGKVGSKCKCHPEAKRMPCAEDYLSVVLNQLCVLHKEKTPVSDRVTKCCTES 480
Db 445 PTLVEVSRNLGKVGSKCKCHPEAKRMPCAEDYLSVVLNQLCVLHKEKTPVSDRVTKCCTES 504

QY 481 LVNRRPCFSALVEDETYVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVVKHKPKAT 540
Db 505 LVNRRPCFSALVEDETYVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVVKHKPKAT 564

QY 541 KEQLKAVMDDFAAFVEKCKCADDKTCFAEBEGKKLVAASQAALGL 585
Db 565 KEQLKAVMDDFAAFVEKCKCADDKTCFAEBEGKKLVAASQAALGL 609
```

```
RESULT 6
US-08-797-689-2
; Sequence 2, Application US/08797689
; Patent No. 5876969
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; APPLICANT: Fournier, Alain
; APPLICANT: Guilton, Jean-Dominique
; APPLICANT: Jung, Gerard
; APPLICANT: Yeh, Patrice
```

;; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,  
;; TITLE OF INVENTION: PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION  
;; TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES  
;; NUMBER OF SEQUENCES: 36  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Rhone-Poulenc Rorer Inc.  
;; STREET: 500 Arcola Road, 3C43  
;; CITY: Collegeville  
;; STATE: PA  
;; COUNTRY: USA  
;; ZIP: 19426  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: Macintosh  
;; OPERATING SYSTEM: System 7.1  
;; SOFTWARE: Word 5.1 (PatentIn)  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/797,689  
;; FILING DATE: 31-JAN-1997  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/256,927  
;; FILING DATE: 28-JUL-1994  
;; APPLICATION NUMBER: FR 92/01064  
;; FILING DATE: 31-JAN-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: PCT/FR93/00085  
;; FILING DATE: 28-JAN-1993  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Smith Ph.D., Julie K.  
;; REGISTRATION NUMBER: P-38,619  
;; REFERENCE/DOCKET NUMBER: ST92006-US  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (610) 454-3839  
;; TELEFAX: (610) 454-3808  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 610 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-08-797-689-2  
Query Match 100.0%; Score 3103; DB 2; Length 610;  
Best Local Similarity 100.0%; Pred. No. 6.9e-287;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DAHKEVAHRRFKDLGEENFKALVLIATAFOYLQOCPEDHVKLVNEVTEFAKTCVADESAS 60  
DB 25 DAHKEVAHRRFKDLGEENFKALVLIATAFOYLQOCPEDHVKLVNEVTEFAKTCVADESAS 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECFLOHKDDNPNLRLVRPEV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECFLOHKDDNPNLRLVRPEV 144  
QY 121 DVMCTAFHNDNEFTLKKLYEIAARRHPYAPPELLFFFAKRYKAAFTTECCQAADKAACLIP 180  
DB 145 DVMCTAFHNDNEFTLKKLYEIAARRHPYAPPELLFFFAKRYKAAFTTECCQAADKAACLIP 204  
QY 181 KLDELURDEGKASAKORLKCASLOKGERAFKAWAVARLSORPPKAEFAVSKLVTDLTK 240  
DB 205 KLDELURDEGKASAKORLKCASLOKGERAFKAWAVARLSORPPKAEFAVSKLVTDLTK 264  
QY 241 VHTECHGDLLECADRADLAKYICENQDISISKLEKCEKPLLEKSHCIAEVENDEMPA 300  
DB 265 VHTECHGDLLECADRADLAKYICENQDISISKLEKCEKPLLEKSHCIAEVENDEMPA 324  
QY 301 DLPSLAADPVESKDVCKNVAEKDVLGMFLFYEARRRHPDYSVLLLRLLAKTYETTLK 360  
DB 325 DLPSLAADPVESKDVCKNVAEKDVLGMFLFYEARRRHPDYSVLLLRLLAKTYETTLK 384  
QY 361 CAAADPHECYAKVDFEFPKPLVEEPQNLIKQNCLEFQLEQYKFNALLVRYTKKVPQVST 420

DB 385 CAAADPHECYAKVDFEFPKPLVEEPQNLIKQNCLEFQLEQYKFNALLVRYTKKVPQVST 444  
QY 421 PTLVEVSRLNGLKVGSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHKEKTPVSDRVTKCCTES 480  
DB 445 PTLVEVSRLNGLKVGSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHKEKTPVSDRVTKCCTES 504  
QY 481 LVNRRPCPSALEVDETYVPKPEFNAETFTFHADICTLSEKEROIKKOTALVELVKKHKPKAT 540  
DB 505 LVNRRPCPSALEVDETYVPKPEFNAETFTFHADICTLSEKEROIKKOTALVELVKKHKPKAT 564  
QY 541 KEOLKAVMDDFAAFVEKCKKADDDKTCFAEEGKKLVAAASQAALGL 585  
DB 565 KEOLKAVMDDFAAFVEKCKKADDDKTCFAEEGKKLVAAASQAALGL 609  
RESULT 7  
US-09-984-186-2  
; Sequence 2, Application US/09984186  
; Patent No. 6866179  
; GENERAL INFORMATION:  
; APPLICANT: Fleer, Reinhard  
; Fournier, Alain  
; Guitton, Jean-Dominique  
; Jung, Gerard  
; Yeh, Patrice  
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,  
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION  
; CONTAINING SAID POLYPEPTIDES  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Rhone-Poulenc Rorer Inc.  
; STREET: 500 Arcola Road, 3C43  
; CITY: Collegeville  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19426  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Macintosh  
; OPERATING SYSTEM: System 7.1  
; SOFTWARE: Word 5.1 (PatentIn)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/984,186  
; FILING DATE: 29-Oct-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/797,689  
; FILING DATE: 31-JAN-1997  
; APPLICATION NUMBER: US 08/256,927  
; FILING DATE: 28-JUL-1994  
; APPLICATION NUMBER: FR 92/01064  
; FILING DATE: 31-JAN-1992  
; APPLICATION NUMBER: PCT/FR93/00085  
; FILING DATE: 28-JAN-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith Ph.D., Julie K.  
; REGISTRATION NUMBER: P-38,619  
; REFERENCE/DOCKET NUMBER: ST92006-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (610) 454-3839  
; TELEFAX: (610) 454-3808  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 610 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-984-186-2

Query Match 100.0%; Score 3103; DB 4; Length 610;  
Best Local Similarity 100.0%; Pred. No. 6.9e-287;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAKSEVAHRFKDLGEENFKALVLIAPFAQYLOQCPFFEDHVKLVNEVTEFAKTCVADESA 60  
Db 25 DAKSEVAHRFKDLGEENFKALVLIAPFAQYLOQCPFFEDHVKLVNEVTEFAKTCVADESA 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNNPRLVVRPEV 120  
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNNPRLVVRPEV 144  
QY 121 DVMCTAFHDNEBETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTTECCQAADKAACLLP 180  
Db 145 DVMCTAFHDNEBETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTTECCQAADKAACLLP 204  
QY 181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSORFPKAEFAEVS KLVTDLTK 240  
Db 205 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSORFPKAEFAEVS KLVTDLTK 264  
QY 241 VHTCCGHDLLLECADDRLADLAKYICENQDSISSKLKECCCKPILLEKSHCIAEVENDEMPA 300  
Db 265 VHTCCGHDLLLECADDRLADLAKYICENQDSISSKLKECCCKPILLEKSHCIAEVENDEMPA 324  
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVVLLRLAKTYETTLEKC 360  
Db 325 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVVLLRLAKTYETTLEKC 384  
QY 361 CAAADPHECYAKVDFEPKPLVEEPQNLIKONCELFEOQLGEYKFQNALLVRYTKVPQVST 420  
Db 385 CAAADPHECYAKVDFEPKPLVEEPQNLIKONCELFEOQLGEYKFQNALLVRYTKVPQVST 444  
QY 421 PTLVEVSRLGKVGSKCKKHPEAKMPCAEADYLSVNLQCLVLEHKTVPVSRVTKCCTES 480  
Db 445 PTLVEVSRLGKVGSKCKKHPEAKMPCAEADYLSVNLQCLVLEHKTVPVSRVTKCCTES 504  
QY 481 LVNRRPCFSALEVDITYPKFNAETFTFHADICTLSEKERQIKKQATALVELVHKPKAT 540  
Db 505 LVNRRPCFSALEVDITYPKFNAETFTFHADICTLSEKERQIKKQATALVELVHKPKAT 564  
QY 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 585  
Db 565 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 609

RESULT 8  
US-09-949-016-11170  
; Sequence 11170, Application US/03949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CLO01307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11170  
; LENGTH: 622  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-11170

Query Match 100.0%; Score 3103; DB 4; Length 622;  
Best Local Similarity 100.0%; Pred. No. 7.1e-287;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAKSEVAHRFKDLGEENFKALVLIAPFAQYLOQCPFFEDHVKLVNEVTEFAKTCVADESA 60  
|||||

Db 38 DAKSEVAHRFKDLGEENFKALVLIAPFAQYLOQCPFFEDHVKLVNEVTEFAKTCVADESA 97  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNNPRLVVRPEV 120  
Db 98 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNNPRLVVRPEV 157  
QY 121 DVMCTAFHDNEBETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTTECCQAADKAACLLP 180  
Db 158 DVMCTAFHDNEBETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTTECCQAADKAACLLP 217  
QY 181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSORFPKAEFAEVS KLVTDLTK 240  
Db 218 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSORFPKAEFAEVS KLVTDLTK 277  
QY 241 VHTCCGHDLLLECADDRLADLAKYICENQDSISSKLKECCCKPILLEKSHCIAEVENDEMPA 300  
Db 278 VHTCCGHDLLLECADDRLADLAKYICENQDSISSKLKECCCKPILLEKSHCIAEVENDEMPA 337  
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVVLLRLAKTYETTLEKC 360  
Db 338 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVVLLRLAKTYETTLEKC 397  
QY 361 CAAADPHECYAKVDFEPKPLVEEPQNLIKONCELFEOQLGEYKFQNALLVRYTKVPQVST 420  
Db 398 CAAADPHECYAKVDFEPKPLVEEPQNLIKONCELFEOQLGEYKFQNALLVRYTKVPQVST 457  
QY 421 PTLVEVSRLGKVGSKCKKHPEAKMPCAEADYLSVNLQCLVLEHKTVPVSRVTKCCTES 480  
Db 458 PTLVEVSRLGKVGSKCKKHPEAKMPCAEADYLSVNLQCLVLEHKTVPVSRVTKCCTES 517  
QY 481 LVNRRPCFSALEVDITYPKFNAETFTFHADICTLSEKERQIKKQATALVELVHKPKAT 540  
Db 518 LVNRRPCFSALEVDITYPKFNAETFTFHADICTLSEKERQIKKQATALVELVHKPKAT 577  
QY 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 585  
Db 578 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 622

RESULT 9  
US-08-256-938-2  
; Sequence 2, Application US/08256938  
; Patent No. 5665863  
; GENERAL INFORMATION:  
; APPLICANT: Yeh, Patrice  
; TITLE OF INVENTION: NEW POLYPEPTIDES HAVING GRANULOCYTE  
; COLONY STIMULATING ACTIVITY, PREPARATION THEREOF AND  
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING SAID POLYPEPTIDES  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Rhone-Poulenc Rorer Inc.  
; STREET: 500 Arcola Road, 3C43  
; CITY: Collegeville  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19426  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Macintosh  
; OPERATING SYSTEM: System 7.1  
; SOFTWARE: Word 5.0 (PatentIn)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/256,938  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR 92/01065  
; FILING DATE: 31-JAN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Goodman, Rosanne  
; REGISTRATION NUMBER: 32,534  
; REFERENCE/DOCKET NUMBER: ST92007-US  
; TELECOMMUNICATION INFORMATION:

```

; TELEPHONE: (610) 454-3817
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 783 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-256-938-2

Query Match 100.0%; Score 3103; DB 1; Length 783;
Best Local Similarity 100.0%; Pred. No. 1e-286;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFKDLGSENFKALVLIAPAYLQOCPEFHVKLVNEVTEFAKTCVADESAB 60
Db 25 DAHKEVAHRFKDLGSENFKALVLIAPAYLQOCPEFHVKLVNEVTEFAKTCVADESAB 84
QY 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQBERNECFLOHKDDNPRLVLRPEV 120
Db 85 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQBERNECFLOHKDDNPRLVLRPEV 144
QY 121 DVMCTAFHNDNEETFLKKLYEIARRHPYFYAPPELLFFAKRYKAAFTCCCAADKAACLIP 180
Db 145 DVMCTAFHNDNEETFLKKLYEIARRHPYFYAPPELLFFAKRYKAAFTCCCAADKAACLIP 204
QY 181 KLDELDECKASSAKQRLKCSLQKFGGERAFKAWAVARLSQRPFAEFAEVSCLVTDLTG 240
Db 205 KLDELDECKASSAKQRLKCSLQKFGGERAFKAWAVARLSQRPFAEFAEVSCLVTDLTG 264
QY 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKLKECEKPELLEKSHCIAEVENDEMPA 300
Db 265 VHTCCHGDLLECCADRADLAKYICENQDSISSKLKECEKPELLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVBSKDVCKNYAEAKOVFLGMFLYFYARRHPDYSVLLLRLLAKTYETTLK 360
Db 325 DLPSLAADFVBSKDVCKNYAEAKOVFLGMFLYFYARRHPDYSVLLLRLLAKTYETTLK 384
QY 361 CAADPHECYAKVDFEPLVEEPQNLIKQNCLEFEQLGEYKFONALLVRYTKVPQVST 420
Db 385 CAADPHECYAKVDFEPLVEEPQNLIKQNCLEFEQLGEYKFONALLVRYTKVPQVST 444
QY 421 PTLVEVSRLGKVGSKCKKHPKAMPKCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 445 PTLVEVSRLGKVGSKCKKHPKAMPKCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVHKPKAT 540
Db 505 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVHKPKAT 564
QY 541 KEQLKAVMDDFAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 585
Db 565 KEQLKAVMDDFAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 609

RESULT 10
US-08-256-938-4
; Sequence 4, Application US/08256938
; Patent No. 5665863
; GENERAL INFORMATION:
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NEW POLYPEPTIDES HAVING GRANULOCYTE
; TITLE OF INVENTION: COLONY STIMULATING ACTIVITY, PREPARATION THEREOF AND
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:

```

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.0 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,938
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 92/01065
; FILING DATE: 31-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Goodman, Rosanne
; REGISTRATION NUMBER: 32,534
; REFERENCE/DOCKET NUMBER: ST92007-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3817
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 787 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-256-938-4

Query Match 100.0%; Score 3103; DB 1; Length 787;
Best Local Similarity 100.0%; Pred. No. 1e-286;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFKDLGSENFKALVLIAPAYLQOCPEFHVKLVNEVTEFAKTCVADESAB 60
Db 203 DAHKEVAHRFKDLGSENFKALVLIAPAYLQOCPEFHVKLVNEVTEFAKTCVADESAB 262
QY 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQBERNECFLOHKDDNPRLVLRPEV 120
Db 263 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQBERNECFLOHKDDNPRLVLRPEV 322
QY 121 DVMCTAFHNDNEETFLKKLYEIARRHPYFYAPPELLFFAKRYKAAFTCCCAADKAACLIP 180
Db 323 DVMCTAFHNDNEETFLKKLYEIARRHPYFYAPPELLFFAKRYKAAFTCCCAADKAACLIP 382
QY 181 KLDELDECKASSAKQRLKCSLQKFGGERAFKAWAVARLSQRPFAEFAEVSCLVTDLTG 240
Db 383 KLDELDECKASSAKQRLKCSLQKFGGERAFKAWAVARLSQRPFAEFAEVSCLVTDLTG 442
QY 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKLKECEKPELLEKSHCIAEVENDEMPA 300
Db 443 VHTCCHGDLLECCADRADLAKYICENQDSISSKLKECEKPELLEKSHCIAEVENDEMPA 502
QY 301 DLPSLAADFVBSKDVCKNYAEAKOVFLGMFLYFYARRHPDYSVLLLRLLAKTYETTLK 360
Db 503 DLPSLAADFVBSKDVCKNYAEAKOVFLGMFLYFYARRHPDYSVLLLRLLAKTYETTLK 562
QY 361 CAADPHECYAKVDFEPLVEEPQNLIKQNCLEFEQLGEYKFONALLVRYTKVPQVST 420
Db 563 CAADPHECYAKVDFEPLVEEPQNLIKQNCLEFEQLGEYKFONALLVRYTKVPQVST 622
QY 421 PTLVEVSRLGKVGSKCKKHPKAMPKCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 623 PTLVEVSRLGKVGSKCKKHPKAMPKCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 682
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVHKPKAT 540
Db 683 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVHKPKAT 742
QY 541 KEQLKAVMDDFAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 585
Db 743 KEQLKAVMDDFAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 787

RESULT 11
US-08-797-689-16

```



```
; Sequence 16, Application US/08797689
; Patent No. 5876969
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; APPLICANT: Fournier, Alain
; APPLICANT: Guitton, Jean-Dominique
; APPLICANT: Jung, Gerard
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES.
; TITLE OF INVENTION: PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcolia Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 787 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-797-689-16

Query Match 100.0%; Score 3103; DB 2; Length 787;
Best Local Similarity 100.0%; Pred. No. le-286;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHKEVAHRFDLGEENFKALVILAFAYLQOCPEFDHVKLVNNEVTEFAKTCVADESAAE 60
Db 203 DAHKEVAHRFDLGEENFKALVILAFAYLQOCPEFDHVKLVNNEVTEFAKTCVADESAAE 262
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPMLPRLVRPEV 120
Db 263 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPMLPRLVRPEV 322
Qy 121 DVMCTAFHDNEETFLKKLYEIARHPHYFYAPELLFFAKRYKAAATECCQAADKAACLLP 180
Db 323 DVMCTAFHDNEETFLKKLYEIARHPHYFYAPELLFFAKRYKAAATECCQAADKAACLLP 382
Qy 181 KLDELURDEGKASSAKORLKASLOKFGRAFKAFAVARLSQFPKAEFAEVSCLKVTDLTG 240
Db 383 KLDELURDEGKASSAKORLKASLOKFGRAFKAFAVARLSQFPKAEFAEVSCLKVTDLTG 442
Qy 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCECKPDLLEKSHCIAEVENDEMPA 300
;
443 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCECKPDLLEKSHCIAEVENDEMPA 502
Qy 301 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYEVARRHPDYSVVLLRLAKTYETTTLEKC 360
Db 503 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYEVARRHPDYSVVLLRLAKTYETTTLEKC 562
Qy 361 CAAADPHECYAKVFDEFKPLVEEPQNLIKONCELFQELGEYKFQNALLVRYTKKVPQVST 420
Db 563 CAAADPHECYAKVFDEFKPLVEEPQNLIKONCELFQELGEYKFQNALLVRYTKKVPQVST 622
Qy 421 PTLVEVSRNLGKVGSKCKKHPEAKMPCAEADYLSVNLNQLCVLHEKTPVSDRVTTCCTES 480
Db 623 PTLVEVSRNLGKVGSKCKKHPEAKMPCAEADYLSVNLNQLCVLHEKTPVSDRVTTCCTES 682
Qy 481 LVNRRPCFSALEVDVETVYPKEFNAETFTFHADICTLSEKEROIKKQTALVELVKKHKPKAT 540
Db 683 LVNRRPCFSALEVDVETVYPKEFNAETFTFHADICTLSEKEROIKKQTALVELVKKHKPKAT 742
Qy 541 KEQLKAVMDDFAAFVEKCKKADDDKTCFAEEGKKLVAASQAALGL 585
Db 743 KEQLKAVMDDFAAFVEKCKKADDDKTCFAEEGKKLVAASQAALGL 787

RESULT 12
US-09-984-186-16
; Sequence 16, Application US/09984186
; Patent No. 6686179
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guitton, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES.
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcolia Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/984,186
; FILING DATE: 29-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 787 amino acids
; TYPE: amino acid
```



```

;
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-09-984-186-16

Query Match 100.0%; Score 3103; DB 4; Length 787;
Best Local Similarity 100.0%; Pred. No. 1e-286;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDGLGEENFKALVLIATAFYAQYLOQCPEFHVHKLNVETFEAKTCVADESAAE 60
Db 203 DAHKSEVAHRFKDGLGEENFKALVLIATAFYAQYLOQCPEFHVHKLNVETFEAKTCVADESAAE 262

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVLRPVP 120
Db 263 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVLRPVP 322

QY 121 DVMCTAFHNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
Db 323 DVMCTAFHNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 382

QY 181 KLDELDRDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 240
Db 383 KLDELDRDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 442

QY 241 VHTTECHGDLLECADRADLAKYICENQDSISSKLKECKEPPLLEKSHCIAEVNDMPA 300
Db 443 VHTTECHGDLLECADRADLAKYICENQDSISSKLKECKEPPLLEKSHCIAEVNDMPA 502

QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLFYEARHPDYSVVLLRLAKTYETTLK 360
Db 503 DLPSLAADFVESKDVCKNYAEAKDVLGMFLFYEARHPDYSVVLLRLAKTYETTLK 562

QY 361 CAAADPHCEYAKVDFBFPLVEEPQNLIKQNCLEFQELGEYKFQNALVRYTKVPQVST 420
Db 563 CAAADPHCEYAKVDFBFPLVEEPQNLIKQNCLEFQELGEYKFQNALVRYTKVPQVST 622

QY 421 PTLVEVSRLNKGKSKCKHPKAPCAEDYLSVVNLQNLVHHEKTPVSDRVTKCCTES 480
Db 623 PTLVEVSRLNKGKSKCKHPKAPCAEDYLSVVNLQNLVHHEKTPVSDRVTKCCTES 682

QY 481 LVNRRPCFSALEVDITYVPKEFNAETFTFHADICTLSEKEROIKKQATALVELVGHKPKAT 540
Db 683 LVNRRPCFSALEVDITYVPKEFNAETFTFHADICTLSEKEROIKKQATALVELVGHKPKAT 742

QY 541 KEQLKAVMDDFAAFVEKCKKADDDKTCFAEKGKLVAAASQAALGL 585
Db 743 KEQLKAVMDDFAAFVEKCKKADDDKTCFAEKGKLVAAASQAALGL 787

RESULT 13
US-08-222-619-3
; Sequence 3, Application US/08222619
; Patent No. 5652352
; GENERAL INFORMATION:
; APPLICANT: Lichenstein, Henri
; APPLICANT: Lyons, David
; APPLICANT: Wurfel, Mark
; APPLICANT: Wright, Samuel
; TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Center, Patent Operations/RRC
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: U.S.
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```

;
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/222,619
; FILING DATE:
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 609 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-222-619-3

Query Match 99.9%; Score 3099; DB 1; Length 609;
Best Local Similarity 99.8%; Pred. No. 1.7e-286;
Matches 584; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDGLGEENFKALVLIATAFYAQYLOQCPEFHVHKLNVETFEAKTCVADESAAE 60
Db 25 DAHKSEVAHRFKDGLGEENFKALVLIATAFYAQYLOQCPEFHVHKLNVETFEAKTCVADESAAE 84

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVLRPVP 120
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVLRPVP 144

QY 121 DVMCTAFHNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
Db 145 DVMCTAFHNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 204

QY 181 KLDELDRDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 240
Db 205 KLDELDRDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 264

QY 241 VHTTECHGDLLECADRADLAKYICENQDSISSKLKECKEPPLLEKSHCIAEVNDMPA 300
Db 265 VHTTECHGDLLECADRADLAKYICENQDSISSKLKECKEPPLLEKSHCIAEVNDMPA 324

QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLFYEARHPDYSVVLLRLAKTYETTLK 360
Db 325 DLPSLAADFVESKDVCKNYAEAKDVLGMFLFYEARHPDYSVVLLRLAKTYETTLK 384

QY 361 CAAADPHCEYAKVDFBFPLVEEPQNLIKQNCLEFQELGEYKFQNALVRYTKVPQVST 420
Db 385 CAAADPHCEYAKVDFBFPLVEEPQNLIKQNCLEFQELGEYKFQNALVRYTKVPQVST 444

QY 421 PTLVEVSRLNKGKSKCKHPKAPCAEDYLSVVNLQNLVHHEKTPVSDRVTKCCTES 480
Db 445 PTLVEVSRLNKGKSKCKHPKAPCAEDYLSVVNLQNLVHHEKTPVSDRVTKCCTES 504

QY 481 LVNRRPCFSALEVDITYVPKEFNAETFTFHADICTLSEKEROIKKQATALVELVGHKPKAT 540
Db 505 LVNRRPCFSALEVDITYVPKEFNAETFTFHADICTLSEKEROIKKQATALVELVGHKPKAT 564

QY 541 KEQLKAVMDDFAAFVEKCKKADDDKTCFAEKGKLVAAASQAALGL 585
Db 565 KEQLKAVMDDFAAFVEKCKKADDDKTCFAEKGKLVAAASQAALGL 609

RESULT 14
US-08-433-037-4
; Sequence 4, Application US/08433037
; Patent No. 5707828
; GENERAL INFORMATION:
; APPLICANT: Sreekrishna, Kotikanyadan
; APPLICANT: Barr, Kathryn A.
; APPLICANT: Briarley, Russell A.
; APPLICANT: Thill, Gregory P.
; APPLICANT: Tschopp, Juerg F.
; TITLE OF INVENTION: EXPRESSION OF HUMAN SERUM ALBUMIN IN
; TITLE OF INVENTION: PICHIA PASTORIS
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
```

; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 11530-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/433,037
; FILING DATE: 03-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9108Z
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 609 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-433-037-4

Query Match 99.9%; Score 3099; DB 1; Length 609;
Best Local Similarity 99.8%; Pred No. 1.7e-286;
Matches 584; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGEENFKALVLI AFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 25 DAHKSEVAHRFKDLGEENFKALVLI AFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 84
QY 61 NCDSLSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 85 NCDSLSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144
QY 121 DVMCTAFHDNBEFTLKKYLYEIARRHPYFYAPELFFAKRYKAAFTCCQAAADKAAACLLP 180
DB 145 DVMCTAFHDNBEFTLKKYLYEIARRHPYFYAPELFFAKRYKAAFTCCQAAADKAAACLLP 204
QY 181 KLDELRLDEGKASSAKQRLKCA SLQKFGERAFKAWAVARLSORFPKAEFAEVS KLVTDLTK 240
DB 205 KLDELRLDEGKVSSAKQRLKCA SLQKFGERAFKAWAVARLSORFPKAEFAEVS KLVTDLTK 264
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 300
DB 265 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYEVARRHPDYSVLLLLRLAKTYETTLTK 360
DB 325 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYEVARRHPDYSVLLLLRLAKTYETTLTK 384
QY 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELFQELGEYKFQNALVRYTKKVPQVST 420
DB 385 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELFQELGEYKFQNALVRYTKKVPQVST 444
QY 421 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 445 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDVETYPKPEFNAETFTFHADICTLSEKERQIKKQTALVELVXHKPKAT 540
DB 505 LVNRRPCFSALEVDVETYPKPEFNAETFTFHADICTLSEKERQIKKQTALVELVXHKPKAT 564
QY 541 KEQLKAVMDDFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 585

; Sequence 3, Application PC/TUS9504075
RESULT 16
PCT-US95-04075-3
; Sequence 3, Application PC/TUS9504075

DB 565 KEQLKAVMDDFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 609
RESULT 15
US-08-897-956A-2
; Sequence 2, Application US/08897956A
; Patent No. 6423512
; GENERAL INFORMATION:
; APPLICANT: Mary Ellen Digan
; APPLICANT: Philip Lake
; APPLICANT: Hermann Gram
; TITLE OF INVENTION: Fusion Polypeptides
; FILE REFERENCE: 600-7244/CPA
; CURRENT APPLICATION NUMBER: US/08/897,956A
; CURRENT FILING DATE: 1997-07-21
; PRIOR APPLICATION NUMBER: 60/022,689
; PRIOR FILING DATE: 1996-07-26
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-08-897-956A-2
Query Match 99.9%; Score 3099; DB 4; Length 609;
Best Local Similarity 99.8%; Pred No. 1.7e-286;
Matches 584; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGEENFKALVLI AFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 25 DAHKSEVAHRFKDLGEENFKALVLI AFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 84
QY 61 NCDSLSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 85 NCDSLSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144
QY 121 DVMCTAFHDNBEFTLKKYLYEIARRHPYFYAPELFFAKRYKAAFTCCQAAADKAAACLLP 180
DB 145 DVMCTAFHDNBEFTLKKYLYEIARRHPYFYAPELFFAKRYKAAFTCCQAAADKAAACLLP 204
QY 181 KLDELRLDEGKASSAKQRLKCA SLQKFGERAFKAWAVARLSORFPKAEFAEVS KLVTDLTK 240
DB 205 KLDELRLDEGKASSAKQRLKCA SLQKFGERAFKAWAVARLSORFPKAEFAEVS KLVTDLTK 264
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 300
DB 265 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYEVARRHPDYSVLLLLRLAKTYETTLTK 360
DB 325 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYEVARRHPDYSVLLLLRLAKTYETTLTK 384
QY 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELFQELGEYKFQNALVRYTKKVPQVST 420
DB 385 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELFQELGEYKFQNALVRYTKKVPQVST 444
QY 421 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 445 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDVETYPKPEFNAETFTFHADICTLSEKERQIKKQTALVELVXHKPKAT 540
DB 505 LVNRRPCFSALEVDVETYPKPEFNAETFTFHADICTLSEKERQIKKQTALVELVXHKPKAT 564
QY 541 KEQLKAVMDDFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 585
DB 565 KEQLKAVMDDFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 609
RESULT 16
PCT-US95-04075-3
; Sequence 3, Application PC/TUS9504075

GENERAL INFORMATION:  
APPLICANT: AMGEN INC.  
TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like  
NUMBER OF INVENTIONS: 33  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Amgen Center, Patent Operations/RRC  
STREET: 1840 DeHavilland Drive  
CITY: Thousand Oaks  
STATE: California  
COUNTRY: U.S.  
ZIP: 91320-1789  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/04075  
FILING DATE:  
CLASSIFICATION:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 609 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
PCT-US95-04075-3

Query Match 99.9%; Score 3099; DB 5; Length 609;  
Best Local Similarity 99.8%; Pred. No. 1.7e-286;  
Matches 584; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DAHKEVAHRFKDGLGENFKALVLIAPQYLOQCPEDHVKLVNEVTEFAKTCVADESAE 60  
DB 25 DAHKEVAHRFKDGLGENFKALVLIAPQYLOQCPEDHVKLVNEVTEFAKTCVADESAE 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVLRPEV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVLRPEV 144  
QY 121 DVMCTAFHNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCECCQAADKAAACLLP 180  
DB 145 DVMCTAFHNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCECCQAADKAAACLLP 204  
QY 181 KLDELDRDEGKASSAKORLKASLOKFGERAFAKAWAVARLSORPPKAEFAEVSKLVTDLTk 240  
DB 205 KLDELDRDEGKASSAKORLKASLOKFGERAFAKAWAVARLSORPPKAEFAEVSKLVTDLTk 264  
QY 241 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLEKCEKPLLEKSHHCIAEVENDEMPA 300  
DB 265 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLEKCEKPLLEKSHHCIAEVENDEMPA 324  
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEYARRHPDYSVWLLRLAKTYETTLK 360  
DB 325 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEYARRHPDYSVWLLRLAKTYETTLK 384  
QY 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELPQOLGEYKFQNALVRYTKVPQVST 420  
DB 385 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELPQOLGEYKFQNALVRYTKVPQVST 444  
QY 421 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 445 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504  
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKQTALVELVKGHPKAT 540  
DB 505 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKQTALVELVKGHPKAT 564  
QY 541 KEOLKAVMDDDFAAFVEKCKCKADDKETCFAEEGKLVAAASQAALGL 585  
DB 565 KEOLKAVMDDDFAAFVEKCKCKADDKETCFAEEGKLVAAASQAALGL 609

## RESULT 17

US-08-897-956A-3  
Sequence 3, Application US/08897956A  
Patent No. 64231512  
GENERAL INFORMATION:  
APPLICANT: Mary Ellen Digan  
APPLICANT: Philip Lake  
APPLICANT: Hermann Gram  
TITLE OF INVENTION: Fusion Polypeptides  
FILE REFERENCE: 600-7244/CPA  
CURRENT APPLICATION NUMBER: US/08/897,956A  
CURRENT FILING DATE: 1997-07-21  
PRIOR APPLICATION NUMBER: 60/022,689  
PRIOR FILING DATE: 1996-07-26  
NUMBER OF SEQ ID NOS: 38  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 978  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Fusion polypeptide  
US-08-897-956A-3

Query Match 99.7%; Score 3095; DB 4; Length 978;  
Best Local Similarity 99.8%; Pred. No. 8e-286;  
Matches 583; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFKDGLGENFKALVLIAPQYLOQCPEDHVKLVNEVTEFAKTCVADESAE 60  
DB 212 DAHKEVAHRFKDGLGENFKALVLIAPQYLOQCPEDHVKLVNEVTEFAKTCVADESAE 271  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVLRPEV 120  
DB 272 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVLRPEV 331  
QY 121 DVMCTAFHNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCECCQAADKAAACLLP 180  
DB 332 DVMCTAFHNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCECCQAADKAAACLLP 391  
QY 181 KLDELDRDEGKASSAKORLKASLOKFGERAFAKAWAVARLSORPPKAEFAEVSKLVTDLTk 240  
DB 392 KLDELDRDEGKASSAKORLKASLOKFGERAFAKAWAVARLSORPPKAEFAEVSKLVTDLTk 451  
QY 241 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLEKCEKPLLEKSHHCIAEVENDEMPA 300  
DB 452 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLEKCEKPLLEKSHHCIAEVENDEMPA 511  
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEYARRHPDYSVWLLRLAKTYETTLK 360  
DB 512 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEYARRHPDYSVWLLRLAKTYETTLK 571  
QY 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELPQOLGEYKFQNALVRYTKVPQVST 420  
DB 572 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELPQOLGEYKFQNALVRYTKVPQVST 631  
QY 421 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 632 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 691  
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKQTALVELVKGHPKAT 540  
DB 692 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKQTALVELVKGHPKAT 751  
QY 541 KEOLKAVMDDDFAAFVEKCKCKADDKETCFAEEGKLVAAASQAALGL 584  
DB 752 KEOLKAVMDDDFAAFVEKCKCKADDKETCFAEEGKLVAAASQAALGL 795

## RESULT 18

US-08-448-196A-3

```
; Sequence 3, Application US/08448196A
; Patent No. 5780594
; GENERAL INFORMATION:
; APPLICANT: CARTER, DANIEL C.
; TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS
; TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR
; TITLE OF INVENTION: RELATED PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NASA
; STREET: MARSHALL SPACE FLIGHT CENTER
; CITY: HUNTSVILLE
; STATE: ALABAMA
; COUNTRY: USA
; ZIP: 35812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448,196A
; FILING DATE: 23-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: BROAD JR., ROBERT L.
; REGISTRATION NUMBER: 18,757
; REFERENCE/DOCKET NUMBER: XX/MFS-28402-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 205-544-0021
; TELEFAX: 205-544-0258
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; US-08-448-196A-3

Query Match 99.7%; Score 3093; DB 1; Length 585;
Best Local Similarity 99.7%; Pred. No. 5.8e-286;
Matches 583; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGGEENFKALVLIAPFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
Db 1 DAHKSEVAHRFKDLGGEENFKALVLIAPFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPMLPRLVRPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPMLPRLVRPEV 120
QY 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP 180
Db 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP 180
QY 181 KLDELREGKASSAKQRLKCSLQKFGERAFAKAWAVARLSORFPKAEFAEYVKLVTDLTJK 240
Db 181 KLDELREGKASSAKQRLKCSLQKFGERAFAKAWAVARLSORFPKAEFAEYVKLVTDLTJK 240
QY 241 VHTECCGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300
Db 241 VHTECCGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYEYARRHPDYSVLLRLAKTYETTTLEKC 360
Db 301 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYEYARRHPDYSVLLRLAKTYETTTLEKC 360
QY 361 CAAADPHCEYAKVDFEFPKPLVEEPQNLIKQNCLEPQLGGEYKFQNALLVRYTKKVPQVST 420
Db 361 CAAADPHCEYAKVDFEFPKPLVEEPQNLIKQNCLEPQLGGEYKFQNALLVRYTKKVPQVST 420
QY 421 PTLVEVSRNLGKVGSKCKKHPKAEKMPKAEADYLSVVLNQLCVLHKEKTPVSDRVTKCCTES 480
Db 421 PTLVEVSRNLGKVGSKCKKHPKAEKMPKAEADYLSVVLNQLCVLHKEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDVETVVPKEFNAETTFHADICTLSEKERQIKKQTALVELVHKHKPKAT 540
Db 481 LVNRRPCFSALEVDVETVVPKEFNAETTFHADICTLSEKERQIKKQTALVELVHKHKPKAT 540

RESULT 19
US-08-984-176-1
; Sequence 1, Application US/08984176
; Patent No. 5948609
; GENERAL INFORMATION:
; APPLICANT: CARTER, DANIEL C
; APPLICANT: HO, JOSEPH X
; APPLICANT: RUKER, FLORIAN
; TITLE OF INVENTION: OXYGEN-TRANSPORTING ALBUMIN-BASED BLOOD REPLACEMENT
; TITLE OF INVENTION: COMPOSITION AND BLOOD VOLUME EXPANDER
; FILE REFERENCE: 08/984,176
; CURRENT APPLICATION NUMBER: US/08/984,176
; CURRENT FILING DATE: 1997-12-03
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-08-984-176-1

Query Match 99.7%; Score 3093; DB 2; Length 585;
Best Local Similarity 99.7%; Pred. No. 5.8e-286;
Matches 583; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGGEENFKALVLIAPFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
Db 1 DAHKSEVAHRFKDLGGEENFKALVLIAPFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPMLPRLVRPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPMLPRLVRPEV 120
QY 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP 180
Db 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP 180
QY 181 KLDELREGKASSAKQRLKCSLQKFGERAFAKAWAVARLSORFPKAEFAEYVKLVTDLTJK 240
Db 181 KLDELREGKASSAKQRLKCSLQKFGERAFAKAWAVARLSORFPKAEFAEYVKLVTDLTJK 240
QY 241 VHTECCGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300
Db 241 VHTECCGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYEYARRHPDYSVLLRLAKTYETTTLEKC 360
Db 301 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYEYARRHPDYSVLLRLAKTYETTTLEKC 360
QY 361 CAAADPHCEYAKVDFEFPKPLVEEPQNLIKQNCLEPQLGGEYKFQNALLVRYTKKVPQVST 420
Db 361 CAAADPHCEYAKVDFEFPKPLVEEPQNLIKQNCLEPQLGGEYKFQNALLVRYTKKVPQVST 420
QY 421 PTLVEVSRNLGKVGSKCKKHPKAEKMPKAEADYLSVVLNQLCVLHKEKTPVSDRVTKCCTES 480
Db 421 PTLVEVSRNLGKVGSKCKKHPKAEKMPKAEADYLSVVLNQLCVLHKEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDVETVVPKEFNAETTFHADICTLSEKERQIKKQTALVELVHKHKPKAT 540
Db 481 LVNRRPCFSALEVDVETVVPKEFNAETTFHADICTLSEKERQIKKQTALVELVHKHKPKAT 540
```



**THIS PAGE BLANK (USPTO)**



85 3103 100.0 639 17 US-10-775-180-131 Sequence 131, App  
86 3103 100.0 639 17 US-10-775-180-132 Sequence 132, App  
87 3103 100.0 639 18 US-10-775-204-414 Sequence 414, App  
88 3103 100.0 639 18 US-10-775-204-416 Sequence 416, App  
89 3103 100.0 639 18 US-10-775-204-417 Sequence 417, App  
90 3103 100.0 639 18 US-10-775-204-418 Sequence 418, App  
91 3103 100.0 639 18 US-10-775-204-1285 Sequence 1285, App  
92 3103 100.0 639 18 US-10-775-204-1295 Sequence 1295, App  
93 3103 100.0 640 15 US-10-433-108-15 Sequence 15, App1  
94 3103 100.0 640 15 US-10-433-108-17 Sequence 17, App1  
95 3103 100.0 640 17 US-10-775-180-615 Sequence 615, App  
96 3103 100.0 640 18 US-10-775-204-280 Sequence 280, App  
97 3103 100.0 640 18 US-10-775-204-288 Sequence 288, App  
98 3103 100.0 640 18 US-10-775-204-1602 Sequence 1602, App  
99 3103 100.0 640 18 US-10-775-204-1603 Sequence 1603, App  
100 3103 100.0 640 18 US-10-775-204-1627 Sequence 1627, App

ALIGNMENTS

RESULT 1  
US-09-929-552-2  
; Sequence 2, Application US/09929552  
; Patent No. US20020123080A1  
; GENERAL INFORMATION:  
; APPLICANT: Sonnenschein, Carlos  
; Soto, Ana M.  
; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Medlen & Carroll, LLP  
; STREET: 220 Montgomery Street, Suite 2200  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States of America  
; ZIP: 94104

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/929,552  
; FILING DATE: 14-Aug-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/769,746  
; FILING DATE: 19-DEC-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Carroll, Peter G.  
; REGISTRATION NUMBER: 32,837  
; REFERENCE/DOCKET NUMBER: MBRI-02584  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 705-8410  
; TELEFAX: (415) 397-8338  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 585 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-929-552-2

Query Match 100.0%; Score 3103; DB 9; Length 585;  
Best Local Similarity 100.0%; Pred. No. 6.3e-251;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DAHKSEVAHRFKDILGEENFKALVLIATAFYAQYLOQCPFDHVKLVNEVTEFAKTCVADESAAE 60  
DB 1 DAHKSEVAHRFKDILGEENFKALVLIATAFYAQYLOQCPFDHVKLVNEVTEFAKTCVADESAAE 60

QY 61 NCDKSLHTLFGDKLCTVATLTRETYGEMADCCAKOEPERNECFLOHKDDNPNLRLVRREV 120  
DB 61 NCDKSLHTLFGDKLCTVATLTRETYGEMADCCAKOEPERNECFLOHKDDNPNLRLVRREV 120  
QY 121 DVMCTAFHDNBEETLKKYLYEIARRHPYFYAPELLFFAKRYKAAAFTECCQAADKAAACLLP 180  
DB 121 DVMCTAFHDNBEETLKKYLYEIARRHPYFYAPELLFFAKRYKAAAFTECCQAADKAAACLLP 180  
QY 191 KLDELDEGKASSAKQRLKASLOKFGERAFAKAWARLSORFPAEFAEYVKLVTDLT 240  
DB 181 KLDELDEGKASSAKQRLKASLOKFGERAFAKAWARLSORFPAEFAEYVKLVTDLT 240  
QY 241 VHTECHGDLLECCADDDRADLAKYICENODSISSKLKECCEKPLLEKSHCIAEVENDEMPA 300  
DB 241 VHTECHGDLLECCADDDRADLAKYICENODSISSKLKECCEKPLLEKSHCIAEVENDEMPA 300  
QY 301 DPLSLAADFVBSKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVVLLLLAKTYETTLEKC 360  
DB 301 DPLSLAADFVBSKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVVLLLLAKTYETTLEKC 360  
QY 361 CAAADPHCYAKVDFDEKPLVEEPPNLIKQNCLEFQOLGEYKFKQNALLVRTKKVPOVST 420  
DB 361 CAAADPHCYAKVDFDEKPLVEEPPNLIKQNCLEFQOLGEYKFKQNALLVRTKKVPOVST 420  
QY 421 PTLVEVSRNLKVGSKCKCKHPEAKRMPCAEDYLSVNLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 421 PTLVEVSRNLKVGSKCKCKHPEAKRMPCAEDYLSVNLNQLCVLHEKTPVSDRVTKCCTES 480  
QY 481 LVNRRPCFSALEVDVETVVPKEFNAETFTFHADICTLSEKEROIKKQTALVELVHKPKAT 540  
DB 481 LVNRRPCFSALEVDVETVVPKEFNAETFTFHADICTLSEKEROIKKQTALVELVHKPKAT 540  
QY 541 KEOLKAVMDDFAAEVKECKCKADDDKTCFAEESKGLVAASQAALGL 585  
DB 541 KEOLKAVMDDFAAEVKECKCKADDDKTCFAEESKGLVAASQAALGL 585

RESULT 2

US-09-932-613-445  
; Sequence 445, Application US/09932613  
; Publication No. US20030091565A1  
; GENERAL INFORMATION:  
; APPLICANT: Human Genome Sciences, Inc.  
; APPLICANT: Beltzer, James P.  
; APPLICANT: Potter, M. Daniel  
; APPLICANT: Fleming, Tony J.  
; APPLICANT: Rosen, Craig A.  
; TITLE OF INVENTION: BINDING POLYPEPTIDES AND METHODS BASED THEREON  
; FILE REFERENCE: DYX-025.1 PCT; DYX-025.1 US  
; CURRENT APPLICATION NUMBER: US/09/932,613  
; CURRENT FILING DATE: 2001-08-17  
; NUMBER OF SEQ ID NOS: 458  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 445  
; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: HomoSapiens  
US-09-932-613-445

Query Match 100.0%; Score 3103; DB 10; Length 585;  
Best Local Similarity 100.0%; Pred. No. 6.3e-251;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDILGEENFKALVLIATAFYAQYLOQCPFDHVKLVNEVTEFAKTCVADESAAE 60  
DB 1 DAHKSEVAHRFKDILGEENFKALVLIATAFYAQYLOQCPFDHVKLVNEVTEFAKTCVADESAAE 60  
QY 61 NCDKSLHTLFGDKLCTVATLTRETYGEMADCCAKOEPERNECFLOHKDDNPNLRLVRREV 120  
DB 61 NCDKSLHTLFGDKLCTVATLTRETYGEMADCCAKOEPERNECFLOHKDDNPNLRLVRREV 120  
QY 121 DVMCTAFHDNBEETLKKYLYEIARRHPYFYAPELLFFAKRYKAAAFTECCQAADKAAACLLP 180



Db 121 DVMCTAFHDNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLUP 180  
QY 181 KLDELDEGKASSAKQRLKCAISLOKFGERAFAKAWAVARLSQRPFPKAEFAEVSCLVTDLTk 240  
Db 181 KLDELDEGKASSAKQRLKCAISLOKFGERAFAKAWAVARLSQRPFPKAEFAEVSCLVTDLTk 240  
QY 241 VHTTECHGDLLECADRADLAKYICENQDSISSKLKECCEKPLLEKSHGICIAEVENDEMPA 300  
Db 241 VHTTECHGDLLECADRADLAKYICENQDSISSKLKECCEKPLLEKSHGICIAEVENDEMPA 300  
QY 301 DLPSSLAADPFVESKDVCKNYAEAKDVLGMLFYEARHPDYSVLLLRLLAKTYETTTLEK 360  
Db 301 DLPSSLAADPFVESKDVCKNYAEAKDVLGMLFYEARHPDYSVLLLRLLAKTYETTTLEK 360  
QY 361 CAADPHCEYAKVDFEFKPLVEEPQNLIKQNCLEFQELGEYKFNALLVRYTKVPQVST 420  
Db 361 CAADPHCEYAKVDFEFKPLVEEPQNLIKQNCLEFQELGEYKFNALLVRYTKVPQVST 420  
QY 421 PTLVEVSRLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
Db 421 PTLVEVSRLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
QY 481 LVNRRPCFSALEVDITYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVGHKPKAT 540  
Db 481 LVNRRPCFSALEVDITYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVGHKPKAT 540  
QY 541 KEQLKAVMDDDFAAFVEKCKCKADDKETCFABEGKCLVAASQAALGL 585  
Db 541 KEQLKAVMDDDFAAFVEKCKCKADDKETCFABEGKCLVAASQAALGL 585

RESULT 3

US-09-984-010-26  
; Sequence 26, Application US/09984010  
; Publication No. US20030104578A1  
; GENERAL INFORMATION:  
; APPLICANT: Ballance, David James  
; TITLE OF INVENTION: RECOMBINANT FUSION PROTEINS TO GROWTH HORMONE  
; AND SERUM ALBUMIN  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FINNEGAN, HENDERSON, PARABOW, GARRETT & DUNNER, LLP  
; STREET: 1300 I Street, NW  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005-3315  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (BPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/984,010  
; FILING DATE: 21-May-2002  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/091,873  
; FILING DATE: 25-JUN-1998  
; APPLICATION NUMBER: PCT/GB96/03164  
; FILING DATE: 19-DEC-1996  
; INFORMATION FOR SEQ ID NO: 26:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 585 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: <Unknown>  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; SEQUENCE DESCRIPTION: SEQ ID NO: 26:  
US-09-984-010-26

Query Match 100.0%; Score 3103; DB 10; Length 585;

Best Local Similarity 100.0%; Pred. No. 6.3e-251;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DAHKEVVAHRPDKDLGEENFKALVLIAFAQYLOQCPEFDHVKLVNEVTEFAKTCVADESAE 60  
Db 1 DAHKEVVAHRPDKDLGEENFKALVLIAFAQYLOQCPEFDHVKLVNEVTEFAKTCVADESAE 60  
QY 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECFLOHQDDPNLPRLVRPEV 120  
Db 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECFLOHQDDPNLPRLVRPEV 120  
QY 121 DVMCTAFHDNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLUP 180  
Db 121 DVMCTAFHDNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLUP 180  
QY 181 KLDELDEGKASSAKQRLKCAISLOKFGERAFAKAWAVARLSQRPFPKAEFAEVSCLVTDLTk 240  
Db 181 KLDELDEGKASSAKQRLKCAISLOKFGERAFAKAWAVARLSQRPFPKAEFAEVSCLVTDLTk 240  
QY 241 VHTTECHGDLLECADRADLAKYICENQDSISSKLKECCEKPLLEKSHGICIAEVENDEMPA 300  
Db 241 VHTTECHGDLLECADRADLAKYICENQDSISSKLKECCEKPLLEKSHGICIAEVENDEMPA 300  
QY 301 DLPSSLAADPFVESKDVCKNYAEAKDVLGMLFYEARHPDYSVLLLRLLAKTYETTTLEK 360  
Db 301 DLPSSLAADPFVESKDVCKNYAEAKDVLGMLFYEARHPDYSVLLLRLLAKTYETTTLEK 360  
QY 361 CAADPHCEYAKVDFEFKPLVEEPQNLIKQNCLEFQELGEYKFNALLVRYTKVPQVST 420  
Db 361 CAADPHCEYAKVDFEFKPLVEEPQNLIKQNCLEFQELGEYKFNALLVRYTKVPQVST 420  
QY 421 PTLVEVSRLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
Db 421 PTLVEVSRLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
QY 481 LVNRRPCFSALEVDITYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVGHKPKAT 540  
Db 481 LVNRRPCFSALEVDITYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVGHKPKAT 540  
QY 541 KEQLKAVMDDDFAAFVEKCKCKADDKETCFABEGKCLVAASQAALGL 585  
Db 541 KEQLKAVMDDDFAAFVEKCKCKADDKETCFABEGKCLVAASQAALGL 585

RESULT 4

US-09-833-041-18  
; Sequence 18, Application US/09833041  
; Publication No. US20030125247A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, Craig A.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF545  
; CURRENT APPLICATION NUMBER: US/09/833,041  
; CURRENT FILING DATE: 2001-04-12  
; PRIOR APPLICATION NUMBER: 60/229,358  
; PRIOR FILING DATE: 2000-04-12  
; PRIOR APPLICATION NUMBER: 60/256,931  
; PRIOR FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/199,384  
; PRIOR FILING DATE: 2000-04-25  
; NUMBER OF SEQ ID NOS: 79  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 18  
; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-833-041-18

Query Match 100.0%; Score 3103; DB 10; Length 585;  
Best Local Similarity 100.0%; Pred. No. 6.3e-251;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	DAHSEVAHRFKDOLGEENFKALVLJAFAYQLQOCPPFEDHVKLVNEVTEFAKTCVADESAB	60
Db	1	DAHSEVAHRFKDOLGEENFKALVLJAFAYQLQOCPPFEDHVKLVNEVTEFAKTCVADESAB	60
Qy	61	NCDKSLHTLFGDKLCTVATLRETYGEMADCAKQEPERNECFLOHKDDNPRLVRPVE	120
Db	61	NCDKSLHTLFGDKLCTVATLRETYGEMADCAKQEPERNECFLOHKDDNPRLVRPVE	120
Qy	121	DVMTAFHDNEETFLLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP	180
Db	121	DVMTAFHDNEETFLLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP	180
Qy	181	KLDELDEBKASAKORUKCASLOKFGGRAPKAWAVARLSORPPKABFAEVSKLVTDLTK	240
Db	181	KLDELDEBKASAKORUKCASLOKFGGRAPKAWAVARLSORPPKABFAEVSKLVTDLTK	240
Qy	241	VHTECCHGDLLECADDRADLAKYICENODSISSSKLKCCCKPILLEKSHCIAEVENDEMPA	300
Db	241	VHTECCHGDLLECADDRADLAKYICENODSISSSKLKCCCKPILLEKSHCIAEVENDEMPA	300
Qy	301	DLPSLAADFVESKDVCKNYABAKOVFLGMFLIYEYARRHPDYSVVLLRLAKTYETTLKRC	360
Db	301	DLPSLAADFVESKDVCKNYABAKOVFLGMFLIYEYARRHPDYSVVLLRLAKTYETTLKRC	360
Qy	361	CAAADPHECIYAKVFDEPKPLVEEPONLIKONCELFEOLGGEYKQONALLVRYTKKVPQVST	420
Db	361	CAAADPHECIYAKVFDEPKPLVEEPONLIKONCELFEOLGGEYKQONALLVRYTKKVPQVST	420
Qy	421	PTLVEVSRNLGVGSKCKKHPEAKMPCABDYLSVNLQCLVHKEKTPVSDRVTCKCTES	480
Db	421	PTLVEVSRNLGVGSKCKKHPEAKMPCABDYLSVNLQCLVHKEKTPVSDRVTCKCTES	480
Qy	481	LVNRRPCFSALEVDETYVPKPFNAETFTFHADICTLSEKERQIKKOTALVELVKHKPKAT	540
Db	481	LVNRRPCFSALEVDETYVPKPFNAETFTFHADICTLSEKERQIKKOTALVELVKHKPKAT	540
Qy	541	KEOLKAVMDDFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL	585
Db	541	KEOLKAVMDDFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL	585

RESULT 5  
US-09-833-117-18  
Sequence 18, Application US/09833117  
Publication No. US20030171267A1  
GENERAL INFORMATION:  
APPLICANT: Rosen, Craig A.  
APPLICANT: Sadeghi, Homa  
APPLICANT: Prior, Christopher P.  
APPLICANT: Turner, Andrew J.  
TITLE OF INVENTION: Albumin Fusion Proteins  
FILE REFERENCE: PF543  
CURRENT APPLICATION NUMBER: US/09/833,117  
CURRENT FILING DATE: 2001-04-12  
PRIOR APPLICATION NUMBER: 60/229,358  
PRIOR FILING DATE: 2000-04-12  
PRIOR APPLICATION NUMBER: 60/256,931  
PRIOR FILING DATE: 2000-12-21  
PRIOR APPLICATION NUMBER: 60/199,384  
PRIOR FILING DATE: 2000-04-25  
NUMBER OF SEQ ID NOS: 36  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 18  
LENGTH: 585  
TYPE: PRT  
ORGANISM: Homo Sapiens  
US-09-833-117-18

1	Db	 DABKSEVAHFKDGLGEENFKALVLIAPAQYLQOCFFEDHVKLVNEVTEFANKTCADESAE	60
61	Qy	 NCBKSJLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPMLPLVRREV	120
61	Db	 NCBKSJLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPMLPLVRREV	120
121	Qy	 DVNCTAFHDNBEETFLKKYLYEIIARRHPYFYAPELLFFAKRYKAAFTTECCQAADKAACLLP	180
121	Db	 DVNCTAFHDNBEETFLKKYLYEIIARRHPYFYAPELLFFAKRYKAAFTTECCQAADKAACLLP	180
181	Qy	 KLDELDRDEGKASSAKORLKCASLOKFGERAFKAWAVARLSORFPKABEAFESKLVTDLTK	240
181	Db	 KLDELDRDEGKASSAKORLKCASLOKFGERAFKAWAVARLSORFPKABEAFESKLVTDLTK	240
241	Qy	 VHTECCHGDDLLECADDDRADJAKYICENQDSISSKLECCCEKPLLEKSHCIAEVDENDEMPA	300
241	Db	 VHTECCHGDDLLECADDDRADJAKYICENQDSISSKLECCCEKPLLEKSHCIAEVDENDEMPA	300
301	Qy	 DLPSLAADFVESKDVCNKYAEAKDVFLGMPLYEYARRHPDYSVLLRLAKTYETTLBKC	360
301	Db	 DLPSLAADFVESKDVCNKYAEAKDVFLGMPLYEYARRHPDYSVLLRLAKTYETTLBKC	360
361	Qy	 CAAADPHECYAKVFDEFPKPLVEEBQNLIQNCSELFQELGEYKFNQNALVRVTKKVPQVST	420
361	Db	 CAAADPHECYAKVFDEFPKPLVEEBQNLIQNCSELFQELGEYKFNQNALVRVTKKVPQVST	420
421	Qy	 PTLVEVSRLNGKVGSKCKHPEAKRMPCEADYLSVNLNQLCVLHEKTPVSDRVTKCCCTES	480
421	Db	 PTLVEVSRLNGKVGSKCKHPEAKRMPCEADYLSVNLNQLCVLHEKTPVSDRVTKCCCTES	480
481	Qy	 LVNRRPCFSALEVDDETVPKFEFNAETFTFHADICTLSEKERQIKKQYALVELVHKHKPKAT	540
481	Db	 LVNRRPCFSALEVDDETVPKFEFNAETFTFHADICTLSEKERQIKKQYALVELVHKHKPKAT	540
541	Qy	 KEQIKAYMDDFAAFVEKCECKADDKETCFABEGKKLVAASQAALGL	585
541	Db	 KEQIKAYMDDFAAFVEKCECKADDKETCFABEGKKLVAASQAALGL	585

```

RESULT 6
US-09-932-322-445
; Sequence 445, Application US/09932322
; Publication No. US2003019473A1
; GENERAL INFORMATION:
; APPLICANT: Dyax Corp.
; APPLICANT: Beltzer, James P.
; APPLICANT: Potter, M. Daniel
; APPLICANT: Fleming, Tony J.
; APPLICANT: Ladner, Robert Charles
; TITLE OF INVENTION: BINDING POLYPEPTIDES FOR B LYMPHOCYTE STIMULATOR PROTEIN (BLYS)
; FILE REFERENCE: DYX-018.1 PCT; DYX-018.1 US
; CURRENT APPLICATION NUMBER: US/09/932,322
; CURRENT FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 458
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 445
; LENGTH: 585
; TYPE: PRT
; ORGANISM: HomoSapiens
; US-09-932-322-445

```

QY 121 DVMCTAFHNEETFLKKYLYEYIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180  
DB 121 DVMCTAFHNEETFLKKYLYEYIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180  
QY 181 KLDELDEGKASSAKORLKCASLOKFGERAFAKAWAVARLSORPPKAEFAEVSKLVTDLTK 240  
DB 181 KLDELDEGKASSAKORLKCASLOKFGERAFAKAWAVARLSORPPKAEFAEVSKLVTDLTK 240  
QY 241 VHTCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300  
DB 241 VHTCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300  
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVLLRLAKTYETTLK 360  
DB 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVLLRLAKTYETTLK 360  
QY 361 CAAADPHCYAKVDFEFKPLVEEPQNLIKONCELFEOQLGEYKFNALLVRYTKKVPQVST 420  
DB 361 CAAADPHCYAKVDFEFKPLVEEPQNLIKONCELFEOQLGEYKFNALLVRYTKKVPQVST 420  
QY 421 PTLVEVSRLNGKVGSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480  
DB 421 PTLVEVSRLNGKVGSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480  
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVGHKPKAT 540  
DB 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVGHKPKAT 540  
QY 541 KEOLKAVMDDFAAAFVEKCKKADDDKTCFAEEGKGLVAASQAALGL 585  
DB 541 KEOLKAVMDDFAAAFVEKCKKADDDKTCFAEEGKGLVAASQAALGL 585

RESULT 7

US-09-832-501-18  
; Sequence 18, Application US/09832501  
; Publication No. US20030199043A1  
; GENERAL INFORMATION:  
; APPLICANT: Ballance, David J.  
; APPLICANT: Sleep, Darrell  
; APPLICANT: Turner, Andrew J.  
; APPLICANT: Sadeghi, Homa  
; APPLICANT: Prior, Christopher P.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF542  
; CURRENT APPLICATION NUMBER: US/09/832,501  
; PRIOR FILING DATE: 2001-04-12  
; PRIOR APPLICATION NUMBER: 60/229,358  
; PRIOR FILING DATE: 2000-04-12  
; PRIOR APPLICATION NUMBER: 60/256,931  
; PRIOR FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/199,384  
; PRIOR FILING DATE: 2000-04-25  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 18  
; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-832-501-18

Query Match 100.0%; Score 3103; DB 10; Length 585;  
Best Local Similarity 100.0%; Pred. No. 6.3e-251;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DAHSEVAHRFKDLGSENFKALVLIAPAYLOQCPEHDHVKLVNEVTEFAKTCVADESAE 60  
DB 1 DAHSEVAHRFKDLGSENFKALVLIAPAYLOQCPEHDHVKLVNEVTEFAKTCVADESAE 60  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120

QY 121 DVMCTAFHNEETFLKKYLYEYIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180  
DB 121 DVMCTAFHNEETFLKKYLYEYIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180  
QY 181 KLDELDEGKASSAKORLKCASLOKFGERAFAKAWAVARLSORPPKAEFAEVSKLVTDLTK 240  
DB 181 KLDELDEGKASSAKORLKCASLOKFGERAFAKAWAVARLSORPPKAEFAEVSKLVTDLTK 240  
QY 241 VHTCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300  
DB 241 VHTCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300  
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVLLRLAKTYETTLK 360  
DB 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVLLRLAKTYETTLK 360  
QY 361 CAAADPHCYAKVDFEFKPLVEEPQNLIKONCELFEOQLGEYKFNALLVRYTKKVPQVST 420  
DB 361 CAAADPHCYAKVDFEFKPLVEEPQNLIKONCELFEOQLGEYKFNALLVRYTKKVPQVST 420  
QY 421 PTLVEVSRLNGKVGSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480  
DB 421 PTLVEVSRLNGKVGSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480  
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVGHKPKAT 540  
DB 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVGHKPKAT 540  
QY 541 KEOLKAVMDDFAAAFVEKCKKADDDKTCFAEEGKGLVAASQAALGL 585  
DB 541 KEOLKAVMDDFAAAFVEKCKKADDDKTCFAEEGKGLVAASQAALGL 585

RESULT 8

US-09-833-118-18  
; Sequence 18, Application US/09833118  
; Publication No. US20030219875A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, Craig A.  
; APPLICANT: Haseltine, William A.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF544  
; CURRENT APPLICATION NUMBER: US/09/833,118  
; PRIOR FILING DATE: 2001-04-12  
; PRIOR APPLICATION NUMBER: 60/229,358  
; PRIOR FILING DATE: 2000-04-12  
; PRIOR APPLICATION NUMBER: 60/256,931  
; PRIOR FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/199,384  
; PRIOR FILING DATE: 2000-04-25  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 18  
; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-833-118-18

Query Match 100.0%; Score 3103; DB 10; Length 585;  
Best Local Similarity 100.0%; Pred. No. 6.3e-251;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DAHSEVAHRFKDLGSENFKALVLIAPAYLOQCPEHDHVKLVNEVTEFAKTCVADESAE 60  
DB 1 DAHSEVAHRFKDLGSENFKALVLIAPAYLOQCPEHDHVKLVNEVTEFAKTCVADESAE 60  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
QY 121 DVMCTAFHNEETFLKKYLYEYIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180  
DB 121 DVMCTAFHNEETFLKKYLYEYIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180

Db 121 DVMCTAFHNDNEETFLKKYLIEIARRHPYFYAPELLFFAKRYKAAATECCQAADKAACLLP 180  
QY 181 KLDELURDEGKASSAKQRLKCAASLOKFGGERAFKAWAVARLSORFPKAEFAEYVKLVTDLTk 240  
Db 181 KLDELURDEGKASSAKQRLKCAASLOKFGGERAFKAWAVARLSORFPKAEFAEYVKLVTDLTk 240  
QY 241 VHTCCGHDLLLECADDRLADLAKYICENQDSISSKLKECCERPLLEKSHCIAEVENDEMPA 300  
Db 241 VHTCCGHDLLLECADDRLADLAKYICENQDSISSKLKECCERPLLEKSHCIAEVENDEMPA 300  
QY 301 DLPSLAADFVESKDVCKNYAEAKOVFLGMFLYIEYARRHPDYSVVLLRLAKTYETTLK 360  
Db 301 DLPSLAADFVESKDVCKNYAEAKOVFLGMFLYIEYARRHPDYSVVLLRLAKTYETTLK 360  
QY 361 CAADPHCEYAKVDFEFPKPLVEEPONLIKQNCLEPEQLGEYKFNALLVRYTKVPQVST 420  
Db 361 CAADPHCEYAKVDFEFPKPLVEEPONLIKQNCLEPEQLGEYKFNALLVRYTKVPQVST 420  
QY 421 PTLVEVSRLGKVGSKCKHPKAEKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
Db 421 PTLVEVSRLGKVGSKCKHPKAEKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
QY 481 LVNRRPCFSALEVDITYVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVKHKPKAT 540  
Db 481 LVNRRPCFSALEVDITYVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVKHKPKAT 540  
QY 541 KEQLKAVMDDFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 585  
Db 541 KEQLKAVMDDFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 585

## RESULT 9

US-09-833-245-18

; Sequence 18, Application US/09833245  
; Publication No. US20040010134A1  
; GENERAL INFORMATION:  
; APPLICANT: Human Genome Sciences, Inc.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; CURRENT APPLICATION NUMBER: US/09/833,245  
; CURRENT FILING DATE: 2001-04-12  
; PRIOR APPLICATION NUMBER: 60/229,358  
; PRIOR FILING DATE: 2000-04-12  
; PRIOR APPLICATION NUMBER: 60/256,931  
; PRIOR FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/199,384  
; PRIOR FILING DATE: 2000-04-25  
; NUMBER OF SEQ ID NOS: 2267  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 18  
; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-833-245-18

Query Match 100.0%; Score 3103; DB 11; Length 585;  
Best Local Similarity 100.0%; Pred. No. 6.3e-251;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DAHSEVAHRFKDLGEENFKALVLIAPAYQLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
Db 1 DAHSEVAHRFKDLGEENFKALVLIAPAYQLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
QY 61 NCDSLSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPMLPRIVRPEV 120  
Db 61 NCDSLSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPMLPRIVRPEV 120  
QY 121 DVMCTAFHNDNEETFLKKYLIEIARRHPYFYAPELLFFAKRYKAAATECCQAADKAACLLP 180  
Db 121 DVMCTAFHNDNEETFLKKYLIEIARRHPYFYAPELLFFAKRYKAAATECCQAADKAACLLP 180  
QY 181 KLDELURDEGKASSAKQRLKCAASLOKFGGERAFKAWAVARLSORFPKAEFAEYVKLVTDLTk 240

Db 181 KLDELURDEGKASSAKQRLKCAASLOKFGGERAFKAWAVARLSORFPKAEFAEYVKLVTDLTk 240  
QY 241 VHTCCGHDLLLECADDRLADLAKYICENQDSISSKLKECCERPLLEKSHCIAEVENDEMPA 300  
Db 241 VHTCCGHDLLLECADDRLADLAKYICENQDSISSKLKECCERPLLEKSHCIAEVENDEMPA 300  
QY 301 DLPSLAADFVESKDVCKNYAEAKOVFLGMFLYIEYARRHPDYSVVLLRLAKTYETTLK 360  
Db 301 DLPSLAADFVESKDVCKNYAEAKOVFLGMFLYIEYARRHPDYSVVLLRLAKTYETTLK 360  
QY 361 CAADPHCEYAKVDFEFPKPLVEEPONLIKQNCLEPEQLGEYKFNALLVRYTKVPQVST 420  
Db 361 CAADPHCEYAKVDFEFPKPLVEEPONLIKQNCLEPEQLGEYKFNALLVRYTKVPQVST 420  
QY 421 PTLVEVSRLGKVGSKCKHPKAEKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
Db 421 PTLVEVSRLGKVGSKCKHPKAEKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
QY 481 LVNRRPCFSALEVDITYVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVKHKPKAT 540  
Db 481 LVNRRPCFSALEVDITYVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVKHKPKAT 540  
QY 541 KEQLKAVMDDFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 585  
Db 541 KEQLKAVMDDFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 585

## RESULT 10

US-09-832-929-18  
; Sequence 18, Application US/09832929  
; Publication No. US20040171123A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, Craig A.  
; APPLICANT: Haseltine, William A.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF547  
; CURRENT APPLICATION NUMBER: US/09/832,929  
; CURRENT FILING DATE: 2001-04-12  
; PRIOR APPLICATION NUMBER: 60/229,358  
; PRIOR FILING DATE: 2000-04-12  
; PRIOR APPLICATION NUMBER: 60/256,931  
; PRIOR FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/199,384  
; PRIOR FILING DATE: 2000-04-25  
; NUMBER OF SEQ ID NOS: 72  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 18  
; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-832-929-18

Query Match 100.0%; Score 3103; DB 11; Length 585;  
Best Local Similarity 100.0%; Pred. No. 6.3e-251;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGEENFKALVLIAPAYQLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
Db 1 DAHSEVAHRFKDLGEENFKALVLIAPAYQLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
QY 61 NCDSLSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPMLPRIVRPEV 120  
Db 61 NCDSLSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPMLPRIVRPEV 120  
QY 121 DVMCTAFHNDNEETFLKKYLIEIARRHPYFYAPELLFFAKRYKAAATECCQAADKAACLLP 180  
Db 121 DVMCTAFHNDNEETFLKKYLIEIARRHPYFYAPELLFFAKRYKAAATECCQAADKAACLLP 180  
QY 181 KLDELURDEGKASSAKQRLKCAASLOKFGGERAFKAWAVARLSORFPKAEFAEYVKLVTDLTk 240  
Db 181 KLDELURDEGKASSAKQRLKCAASLOKFGGERAFKAWAVARLSORFPKAEFAEYVKLVTDLTk 240  
QY 241 VHTCCGHDLLLECADDRLADLAKYICENQDSISSKLKECCERPLLEKSHCIAEVENDEMPA 300

Db 241 VHTCCRGDLLECADRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVNDMPA 300  
QY 301 DLPSLAADFVSKDVCCKNAEAKDVLGMLFYEARHPDYSVLLRLAKTYETTLK 360  
Db 301 DLPSLAADFVSKDVCCKNAEAKDVLGMLFYEARHPDYSVLLRLAKTYETTLK 360  
QY 361 CAAADPHECYAKVDFEFKPLVEBPQNLIKQNCFLFQGLGEYKFNALLVRYTKVPQVST 420  
Db 361 CAAADPHECYAKVDFEFKPLVEBPQNLIKQNCFLFQGLGEYKFNALLVRYTKVPQVST 420  
QY 421 PTLVEVSRNLGVSKCKCHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480  
Db 421 PTLVEVSRNLGVSKCKCHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480  
QY 481 LVNRPCFSALEVDETYVPKEFNAETFTPHADICTLSEKEROIKKQATLVELVGHKPKAT 540  
Db 481 LVNRPCFSALEVDETYVPKEFNAETFTPHADICTLSEKEROIKKQATLVELVGHKPKAT 540  
QY 541 KEQLKAVMDDDFAAFVEKCKKADDDKTCFAEEGKGLVAASQAALGL 585  
Db 541 KEQLKAVMDDDFAAFVEKCKKADDDKTCFAEEGKGLVAASQAALGL 585

## RESULT 11

US-10-153-604A-5  
; Sequence 5, Application US/10153604A  
; Publication No. US20030143191A1  
; GENERAL INFORMATION:  
; APPLICANT: Bell et al.  
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins  
; FILE REFERENCE: PF556  
; CURRENT APPLICATION NUMBER: US/10/153,604A  
; PRIOR FILING DATE: 2002-05-24  
; PRIOR APPLICATION NUMBER: 60/293,212  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 137  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 5  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-10-153-604A-5

Query Match 100.0%; Score 3103; DB 14; Length 585;  
Best Local Similarity 100.0%; Pred. No. 6.3e-251;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DAHKEVAHRFKDLGEENFKALVLIATAQYLOQCPFEDHVKLVNEVTEFAKTCVADES 60  
Db 1 DAHKEVAHRFKDLGEENFKALVLIATAQYLOQCPFEDHVKLVNEVTEFAKTCVADES 60  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRVPRV 120  
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRVPRV 120  
QY 121 DVNCTAFHDNEETFLKKLYEIARRHPYFAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180  
Db 121 DVNCTAFHDNEETFLKKLYEIARRHPYFAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180  
QY 181 KLDELURDEGKASSAKQRLKASLOKFGERAFAKAWAVARLSORFPKAEFAEVSKLVTDLTK 240  
Db 181 KLDELURDEGKASSAKQRLKASLOKFGERAFAKAWAVARLSORFPKAEFAEVSKLVTDLTK 240  
QY 241 VHTCCRGDLLECADRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVNDMPA 300  
Db 241 VHTCCRGDLLECADRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVNDMPA 300  
QY 301 DLPSLAADFVSKDVCCKNAEAKDVLGMLFYEARHPDYSVLLRLAKTYETTLK 360  
Db 301 DLPSLAADFVSKDVCCKNAEAKDVLGMLFYEARHPDYSVLLRLAKTYETTLK 360  
QY 361 CAAADPHECYAKVDFEFKPLVEBPQNLIKQNCFLFQGLGEYKFNALLVRYTKVPQVST 420

Db 361 CAAADPHECYAKVDFEFKPLVEBPQNLIKQNCFLFQGLGEYKFNALLVRYTKVPQVST 420  
QY 421 PTLVEVSRNLGVSKCKCHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480  
Db 421 PTLVEVSRNLGVSKCKCHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480  
QY 481 LVNRPCFSALEVDETYVPKEFNAETFTPHADICTLSEKEROIKKQATLVELVGHKPKAT 540  
Db 481 LVNRPCFSALEVDETYVPKEFNAETFTPHADICTLSEKEROIKKQATLVELVGHKPKAT 540  
QY 541 KEQLKAVMDDDFAAFVEKCKKADDDKTCFAEEGKGLVAASQAALGL 585  
Db 541 KEQLKAVMDDDFAAFVEKCKKADDDKTCFAEEGKGLVAASQAALGL 585

## RESULT 12

US-10-319-263-1  
; Sequence 1, Application US/10319263  
; Publication No. US20030180820A1  
; GENERAL INFORMATION:  
; APPLICANT: Bar-Or M.D., David  
; APPLICANT: Lau Ph.D., Edward  
; APPLICANT: Winkler M.D., James V.  
; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and  
; TITLE OF INVENTION: Kits  
; FILE REFERENCE: ISC007  
; CURRENT APPLICATION NUMBER: US/10/319,263  
; PRIOR FILING DATE: 2002-12-13  
; PRIOR APPLICATION NUMBER: 60/115,392  
; PRIOR FILING DATE: 1999-01-11  
; PRIOR APPLICATION NUMBER: 60/102,738  
; PRIOR FILING DATE: 1998-10-02  
; PRIOR APPLICATION NUMBER: 09/165,926  
; PRIOR FILING DATE: 1998-10-02  
; PRIOR APPLICATION NUMBER: 09/165,581  
; PRIOR FILING DATE: 1998-10-02  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-319-263-1

Query Match 100.0%; Score 3103; DB 14; Length 585;  
Best Local Similarity 100.0%; Pred. No. 6.3e-251;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DAHKEVAHRFKDLGEENFKALVLIATAQYLOQCPFEDHVKLVNEVTEFAKTCVADES 60  
Db 1 DAHKEVAHRFKDLGEENFKALVLIATAQYLOQCPFEDHVKLVNEVTEFAKTCVADES 60  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRVPRV 120  
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRVPRV 120  
QY 121 DVNCTAFHDNEETFLKKLYEIARRHPYFAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180  
Db 121 DVNCTAFHDNEETFLKKLYEIARRHPYFAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180  
QY 181 KLDELURDEGKASSAKQRLKASLOKFGERAFAKAWAVARLSORFPKAEFAEVSKLVTDLTK 240  
Db 181 KLDELURDEGKASSAKQRLKASLOKFGERAFAKAWAVARLSORFPKAEFAEVSKLVTDLTK 240  
QY 241 VHTCCRGDLLECADRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVNDMPA 300  
Db 241 VHTCCRGDLLECADRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVNDMPA 300  
QY 301 DLPSLAADFVSKDVCCKNAEAKDVLGMLFYEARHPDYSVLLRLAKTYETTLK 360  
Db 301 DLPSLAADFVSKDVCCKNAEAKDVLGMLFYEARHPDYSVLLRLAKTYETTLK 360

```
QY 361 CAADPHCEYAKVDFEFKPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKKVPQVST 420
|||||
Db 361 CAADPHCEYAKVDFEFKPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKKVPQVST 420
QY 421 PTLVEVSNLKGVSCKCKHPEAKMPCAEADYLSVVLNQLCVLHKEKTPVSDRVTKCCCTES 480
|||||
Db 421 PTLVEVSNLKGVSCKCKHPEAKMPCAEADYLSVVLNQLCVLHKEKTPVSDRVTKCCCTES 480
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKKPKAT 540
|||||
Db 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKKPKAT 540
QY 541 KEQLKAVNMDDFAAVFEKCKCKADDDKTCFAEBEGKKLVAASQAALGL 585
|||||
Db 541 KEQLKAVNMDDFAAVFEKCKCKADDDKTCFAEBEGKKLVAASQAALGL 585

RESULT 13
US-10-319-263-2
; Sequence 2, Application US/10319263
; Publication No. US20030180820A1
; GENERAL INFORMATION:
; APPLICANT: Bar-Or M.D., David
; APPLICANT: Lau Ph.D., Edward
; APPLICANT: Winkler M.D., James V.
; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and
; TITLE OF INVENTION: Kits
; FILE REFERENCE: ISCO07
; CURRENT APPLICATION NUMBER: US/10/319,263
; CURRENT FILING DATE: 2002-12-13
; PRIOR APPLICATION NUMBER: 60/115,392
; PRIOR FILING DATE: 1999-01-11
; PRIOR APPLICATION NUMBER: 60/102,738
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,926
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,581
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (1)..(585)
; OTHER INFORMATION: ACETYLTATION
US-10-319-263-2

Query Match 100.0%; Score 3103; DB 14; Length 585;
Best Local Similarity 100.0%; Pred. No. 6.3e-251;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRPKDLGEENFKALVLIAFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
Db 1 DAHKSEVAHRPKDLGEENFKALVLIAFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLPLRLVRPEV 120
Db 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLPLRLVRPEV 120
QY 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP 180
Db 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP 180
QY 241 VHTCECHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
Db 241 VHTCECHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
```

```
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFILGMFLYEVARRHPDYSVLLLRLLAKTYETTLK 360
|||||
Db 301 DLPSLAADFVESKDVCKNYAEAKDVFILGMFLYEVARRHPDYSVLLLRLLAKTYETTLK 360
QY 361 CAADPHCEYAKVDFEFKPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKKVPQVST 420
|||||
Db 361 CAADPHCEYAKVDFEFKPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKKVPQVST 420
QY 421 PTLVEVSNLKGVSCKCKHPEAKMPCAEADYLSVVLNQLCVLHKEKTPVSDRVTKCCCTES 480
|||||
Db 421 PTLVEVSNLKGVSCKCKHPEAKMPCAEADYLSVVLNQLCVLHKEKTPVSDRVTKCCCTES 480
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKKPKAT 540
|||||
Db 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKKPKAT 540
QY 541 KEQLKAVNMDDFAAVFEKCKCKADDDKTCFAEBEGKKLVAASQAALGL 585
|||||
Db 541 KEQLKAVNMDDFAAVFEKCKCKADDDKTCFAEBEGKKLVAASQAALGL 585

RESULT 14
US-10-414-469-1
; Sequence 1, Application US/10414469
; Publication No. US20030190691A1
; GENERAL INFORMATION:
; APPLICANT: Bar-Or M.D., David
; APPLICANT: Lau Ph.D., Edward
; APPLICANT: Winkler M.D., James V.
; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and
; TITLE OF INVENTION: Kits
; FILE REFERENCE: ISCO07
; CURRENT APPLICATION NUMBER: US/10/414,469
; CURRENT FILING DATE: 2003-04-15
; PRIOR APPLICATION NUMBER: 09/806,247
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: PCT/US99/22905
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: 60/115,392
; PRIOR FILING DATE: 1999-01-11
; PRIOR APPLICATION NUMBER: 60/102,738
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,926
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,581
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-414-469-1

Query Match 100.0%; Score 3103; DB 14; Length 585;
Best Local Similarity 100.0%; Pred. No. 6.3e-251;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRPKDLGEENFKALVLIAFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
Db 1 DAHKSEVAHRPKDLGEENFKALVLIAFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLPLRLVRPEV 120
Db 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLPLRLVRPEV 120
QY 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP 180
Db 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP 180
QY 181 KLDELREDEGKASSAKQRLKCAVLSQKFGERAFKAWAVARLSQRPKAEFAEVSCLVTDLT 240
```

```
Db 181 KLDELDEGKASSAKORLKCASLOKGERAFKAWAVARLSORPPKAEPAEVSKLVTDLTk 240
Qy 241 VHTTECHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHGICIAEVENDEMPA 300
Db 241 VHTTECHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHGICIAEVENDEMPA 300
Qy 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEARRHDPDYSVVLNLRRAKTYETTTLEK 360
Db 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEARRHDPDYSVVLNLRRAKTYETTTLEK 360
Qy 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCLEFQGLGEYKFQNALLVRYTKVPQVST 420
Db 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCLEFQGLGEYKFQNALLVRYTKVPQVST 420
Qy 421 PTLVEVSRNLGVSKCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 421 PTLVEVSRNLGVSKCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Qy 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVKGHPKAT 540
Db 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVKGHPKAT 540
Qy 541 KEQLKAVMDDDFAAFVEKCKCKADDDKTCFAEEGKKLVAASQAALGL 585
Db 541 KEQLKAVMDDDFAAFVEKCKCKADDDKTCFAEEGKKLVAASQAALGL 585

RESULT 15
US-10-414-469-2
; Sequence 2, Application US/10414469
; Publication No. US20030190691A1
; GENERAL INFORMATION:
; APPLICANT: Bar-Or M.D., David
; APPLICANT: Lau Ph.D., Edward
; APPLICANT: Winkler M.D., James V.
; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and
; TITLE OF INVENTION: Kits
; FILE REFERENCE: ISCO07
; CURRENT APPLICATION NUMBER: US/10/414,469
; CURRENT FILING DATE: 2003-04-15
; PRIOR APPLICATION NUMBER: 09/806,247
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: PCT/US99/22905
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: 60/115,392
; PRIOR FILING DATE: 1999-01-11
; PRIOR APPLICATION NUMBER: 60/102,738
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,926
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,581
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (1)..(585)
; OTHER INFORMATION: ACETYLYATION
US-10-414-469-2
```

```
Query Match 100.0%; Score 3103; DB 14; Length 585;
Best Local Similarity 100.0%; Pred. No. 6.3e-251;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHRFKDGLGENFKALVLIAPQYLOQCPEPDHVKLVNEVTEFAKTCVADESAB 60
Db 1 DAHSEVAHRFKDGLGENFKALVLIAPQYLOQCPEPDHVKLVNEVTEFAKTCVADESAB 60
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECFLOHKDDNPNLRLVRPEV 120
```

```
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECFLOHKDDNPNLRLVRPEV 120
Qy 121 DVMCTAFHDNEETFLKKVLYEYARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAAACLLP 180
Db 121 DVMCTAFHDNEETFLKKVLYEYARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAAACLLP 180
Qy 181 KLDELDEGKASSAKORLKCASLOKGERAFKAWAVARLSORPPKAEPAEVSKLVTDLTk 240
Db 181 KLDELDEGKASSAKORLKCASLOKGERAFKAWAVARLSORPPKAEPAEVSKLVTDLTk 240
Qy 241 VHTTECHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHGICIAEVENDEMPA 300
Db 241 VHTTECHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHGICIAEVENDEMPA 300
Qy 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEARRHDPDYSVVLNLRRAKTYETTTLEK 360
Db 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEARRHDPDYSVVLNLRRAKTYETTTLEK 360
Qy 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCLEFQGLGEYKFQNALLVRYTKVPQVST 420
Db 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCLEFQGLGEYKFQNALLVRYTKVPQVST 420
Qy 421 PTLVEVSRNLGVSKCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 421 PTLVEVSRNLGVSKCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Qy 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVKGHPKAT 540
Db 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVKGHPKAT 540
Qy 541 KEQLKAVMDDDFAAFVEKCKCKADDDKTCFAEEGKKLVAASQAALGL 585
Db 541 KEQLKAVMDDDFAAFVEKCKCKADDDKTCFAEEGKKLVAASQAALGL 585

RESULT 16
US-10-413-831-1
; Sequence 1, Application US/10413831
; Publication No. US20030194813A1
; GENERAL INFORMATION:
; APPLICANT: Bar-Or M.D., David
; APPLICANT: Lau Ph.D., Edward
; APPLICANT: Winkler M.D., James V.
; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and
; TITLE OF INVENTION: Kits
; FILE REFERENCE: ISCO07
; CURRENT APPLICATION NUMBER: US/10/413,831
; CURRENT FILING DATE: 2003-04-15
; PRIOR APPLICATION NUMBER: US/09/806,247
; PRIOR FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 60/115,392
; PRIOR FILING DATE: 1999-01-11
; PRIOR APPLICATION NUMBER: 60/102,738
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,926
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,581
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-413-831-1

Query Match 100.0%; Score 3103; DB 14; Length 585;
Best Local Similarity 100.0%; Pred. No. 6.3e-251;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHRFKDGLGENFKALVLIAPQYLOQCPEPDHVKLVNEVTEFAKTCVADESAB 60
```



Db 1 DAHSEVAHRFKDLGEEFNFKALVLIATAQYLQOCPFEDHVKLVNEVTEFAKTCVADSAE 60  
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLRLVRPEV 120  
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLRLVRPEV 120  
Qy 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
Db 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
Qy 181 KLDELDRDEGKASSAKQRLKCAASLQKFGERAFKAWAVARLSQRFPAEFAEYVKLVTDLTk 240  
Db 181 KLDELDRDEGKASSAKQRLKCAASLQKFGERAFKAWAVARLSQRFPAEFAEYVKLVTDLTk 240  
Qy 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 300  
Db 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 300  
Qy 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEYARRHPDYSVLLLLRLAKTYETTLEKc 360  
Db 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEYARRHPDYSVLLLLRLAKTYETTLEKc 360  
Qy 361 CAADPHECYAKVDFEFKPLVEEPONLIKQNCLEFQOLGEYKFNQALLVRYTKVPQVST 420  
Db 361 CAADPHECYAKVDFEFKPLVEEPONLIKQNCLEFQOLGEYKFNQALLVRYTKVPQVST 420  
Qy 421 PTLVEVSRLGKVGSKCKKHPKAEKMPCAEDYLSVNLQCLVHLHEKTPVSDRVTKCCTES 480  
Db 421 PTLVEVSRLGKVGSKCKKHPKAEKMPCAEDYLSVNLQCLVHLHEKTPVSDRVTKCCTES 480  
Qy 481 LVNRRPCFSALEVDETYVPKEFNAETTFHADICTLSEKERQIKKQATALVELVKHKPKAT 540  
Db 481 LVNRRPCFSALEVDETYVPKEFNAETTFHADICTLSEKERQIKKQATALVELVKHKPKAT 540  
Qy 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 585  
Db 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 585

RESULT 17

US-10-413-831-2  
; Sequence 2, Application US/10413831  
; Publication No. US20030194813A1  
; GENERAL INFORMATION:  
; APPLICANT: Bar-Or M.D., David  
; APPLICANT: Lau Ph.D., Edward  
; APPLICANT: Winkler M.D., James V.  
; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and  
; TITLE OF INVENTION: Kits  
; FILE REFERENCE: ISC007  
; CURRENT APPLICATION NUMBER: US/10/413,831  
; CURRENT FILING DATE: 2003-04-15  
; PRIOR APPLICATION NUMBER: US/09/806,247  
; PRIOR FILING DATE: 2001-07-16  
; PRIOR APPLICATION NUMBER: 60/115,392  
; PRIOR FILING DATE: 1999-01-11  
; PRIOR APPLICATION NUMBER: 60/102,738  
; PRIOR FILING DATE: 1998-10-02  
; PRIOR APPLICATION NUMBER: 09/165,926  
; PRIOR FILING DATE: 1998-10-02  
; PRIOR APPLICATION NUMBER: 09/165,581  
; PRIOR FILING DATE: 1998-10-02  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: Patentin ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: MOD\_RES  
; LOCATION: (1)..(585)  
; OTHER INFORMATION: ACETYLYATION  
US-10-413-831-2

Query Match 100.0%; Score 3103; DB 14; Length 585;  
Best Local Similarity 100.0%; Pred. No. 6.3e-251; Mismatches 0; Indels 0; Gaps 0;  
Matches 585; Conservative 0;  
Qy 1 DAHSEVAHRFKDLGEEFNFKALVLIATAQYLQOCPFEDHVKLVNEVTEFAKTCVADSAE 60  
Db 1 DAHSEVAHRFKDLGEEFNFKALVLIATAQYLQOCPFEDHVKLVNEVTEFAKTCVADSAE 60  
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLRLVRPEV 120  
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLRLVRPEV 120  
Qy 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
Db 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
Qy 181 KLDELDRDEGKASSAKQRLKCAASLQKFGERAFKAWAVARLSQRFPAEFAEYVKLVTDLTk 240  
Db 181 KLDELDRDEGKASSAKQRLKCAASLQKFGERAFKAWAVARLSQRFPAEFAEYVKLVTDLTk 240  
Qy 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 300  
Db 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 300  
Qy 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEYARRHPDYSVLLLLRLAKTYETTLEKc 360  
Db 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEYARRHPDYSVLLLLRLAKTYETTLEKc 360  
Qy 361 CAADPHECYAKVDFEFKPLVEEPONLIKQNCLEFQOLGEYKFNQALLVRYTKVPQVST 420  
Db 361 CAADPHECYAKVDFEFKPLVEEPONLIKQNCLEFQOLGEYKFNQALLVRYTKVPQVST 420  
Qy 421 PTLVEVSRLGKVGSKCKKHPKAEKMPCAEDYLSVNLQCLVHLHEKTPVSDRVTKCCTES 480  
Db 421 PTLVEVSRLGKVGSKCKKHPKAEKMPCAEDYLSVNLQCLVHLHEKTPVSDRVTKCCTES 480  
Qy 481 LVNRRPCFSALEVDETYVPKEFNAETTFHADICTLSEKERQIKKQATALVELVKHKPKAT 540  
Db 481 LVNRRPCFSALEVDETYVPKEFNAETTFHADICTLSEKERQIKKQATALVELVKHKPKAT 540  
Qy 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 585  
Db 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 585

RESULT 18

US-10-413-832-1  
; Sequence 1, Application US/10413832  
; Publication No. US20030215359A1  
; GENERAL INFORMATION:  
; APPLICANT: Bar-Or M.D., David  
; APPLICANT: Lau Ph.D., Edward  
; APPLICANT: Winkler M.D., James V.  
; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and  
; TITLE OF INVENTION: Kits  
; FILE REFERENCE: ISC007  
; CURRENT APPLICATION NUMBER: US/10/413,832  
; CURRENT FILING DATE: 2003-04-15  
; PRIOR APPLICATION NUMBER: US/09/806,247  
; PRIOR FILING DATE: 2001-07-16  
; PRIOR APPLICATION NUMBER: 60/115,392  
; PRIOR FILING DATE: 1999-01-11  
; PRIOR APPLICATION NUMBER: 60/102,738  
; PRIOR FILING DATE: 1998-10-02  
; PRIOR APPLICATION NUMBER: 09/165,926  
; PRIOR FILING DATE: 1998-10-02  
; PRIOR APPLICATION NUMBER: 09/165,581  
; PRIOR FILING DATE: 1998-10-02  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: Patentin ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 585





```
; PRIOR APPLICATION NUMBER: 60/102,738
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,926
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,581
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-414-386-1

Query Match      100.0%; Score 3103; DB 15; Length 585;
Best Local Similarity 100.0%; Pred. No. 6.3e-251;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHKEVAHRFKDLGEENFKALVIAFAQYIQQCPFFEDHVKLVNEVTEFAKTCVADESAAE 60
Db 1 DAHKEVAHRFKDLGEENFKALVIAFAQYIQQCPFFEDHVKLVNEVTEFAKTCVADESAAE 60

Qy 61 NCDKSLHTLFGDKLCTVATLTRETYGEMADCCAKQEPERNECFLOHKDDNPMLPRLVRPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLTRETYGEMADCCAKQEPERNECFLOHKDDNPMLPRLVRPEV 120

Qy 121 DVMCTAFHDNEETFLKKVLYEIARHPYFYAPPELLFFAKRYKAAFTCCQADKAAACLLP 180
Db 121 DVMCTAFHDNEETFLKKVLYEIARHPYFYAPPELLFFAKRYKAAFTCCQADKAAACLLP 180

Qy 181 KLDELDEGKASSAKORLKASLOKFGERAFAKAWARLSORFPAEFAEVSKLVTDLTK 240
Db 181 KLDELDEGKASSAKORLKASLOKFGERAFAKAWARLSORFPAEFAEVSKLVTDLTK 240

Qy 241 VHTCCHGDLLECADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 300
Db 241 VHTCCHGDLLECADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 300

Qy 301 DLPSLAADFVSKDCKNYAEAKDVFGLGMFLYEVARRHPDYSVLLRLAKTYETTLTK 360
Db 301 DLPSLAADFVSKDCKNYAEAKDVFGLGMFLYEVARRHPDYSVLLRLAKTYETTLTK 360

Qy 361 CAADPHECYAKVDFEPKPLVEEPQNLIKQNCLEFQELGEYKFNALLVRYTKKVPQVST 420
Db 361 CAADPHECYAKVDFEPKPLVEEPQNLIKQNCLEFQELGEYKFNALLVRYTKKVPQVST 420

Qy 421 PTLVEVSRNLGKVGSKCKHPKAKMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 421 PTLVEVSRNLGKVGSKCKHPKAKMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480

Qy 481 LVNRRPCFSALEVDVETYPKEFNAETFTFHADICTLSEKERQIKKOTALVELVHKPKAT 540
Db 481 LVNRRPCFSALEVDVETYPKEFNAETFTFHADICTLSEKERQIKKOTALVELVHKPKAT 540

Qy 541 KEQLKAVNMDDFAAVFVEKCKCKADDDKTCFAEGKGLVAASQAALGL 585
Db 541 KEQLKAVNMDDFAAVFVEKCKCKADDDKTCFAEGKGLVAASQAALGL 585
```

Search completed: October 13, 2005, 08:53:52  
Job time : 171 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2005, 08:40:01 ; Search time 43 Seconds

(without alignments)  
1308.995 Million cell updates/sec

Title: US-10-816-042-18

Perfect score: 3103

Sequence: 1 DAHKSEVAHRFKDLGEENFK.....TCFAEGKKLVAASQAALGL 585

Scoring table: BLOSUM62

Gapop: 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

1: PIR 79:\*

2: PIR1:\*

3: PIR2:\*

4: PIR3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3103	100.0	609	1 ABHUS	serum albumin prec
2	2942	94.8	600	2 A47391	serum albumin prec
3	2620	84.4	608	2 S57632	serum albumin prec
4	2475.5	79.8	607	1 ABHOS	serum albumin prec
5	2446.5	78.8	607	1 ABHOS	serum albumin prec
6	2432.5	78.4	607	1 ABHOS	serum albumin prec
7	2426	78.2	608	1 ABHOS	serum albumin prec
8	2411.5	77.7	605	1 ABHOS	serum albumin prec
9	2387	76.9	609	2 JCS838	albumin - Mongolia
10	1861	60.0	453	2 A05139	serum albumin prec
11	1557.5	50.2	615	1 ABCHS	serum albumin prec
12	1253.5	40.4	609	2 JCS258	alpha-fetoprotein
13	1249.5	40.3	609	1 FPHU	alpha-fetoprotein
14	1242.5	40.0	609	1 FPGU	alpha-fetoprotein
15	1205	38.8	607	1 ABXL72	74K albumin prec
16	1181.5	38.1	265	2 I46986	albumin - dog (fra
17	1175.5	37.9	608	1 ABXL68	68K serum albumin
18	1084	34.9	605	1 FPM5	alpha-fetoprotein
19	1067	34.4	611	1 FPR2	alpha-fetoprotein
20	1055	34.0	599	1 A54906	afamin precursor
21	928.5	29.9	614	2 S59517	serum albumin prec
22	928	29.9	608	2 A53195	afamin precursor
23	747.5	24.1	608	1 ABONS1	serum albumin 1 pr
24	742.5	23.9	608	1 ABONS2	serum albumin 2 pr
25	699	22.5	382	2 A37253	serum albumin - bu
26	440.5	14.2	1423	1 S27941	serum albumin - se
27	386	12.4	474	1 VYHDD	vitamin D-binding
28	385	12.4	476	1 VYHDD	vitamin D-binding
29	372	12.0	472	1 A35327	vitamin D-binding

30	184	5.9	1819	2	A71928	cag island protein
31	184	5.9	1827	2	G64585	cag pathogenicity
32	141.5	4.6	1560	2	T30282	calcium-binding pr
33	137.5	4.4	1348	2	AG2558	hypothetical prote
34	134.5	4.3	1004	2	JC2221	major surface glyco
35	134	4.3	1083	2	JC2300	cell surface glyco
36	134	4.3	1780	2	T17272	hypothetical prote
37	133.5	4.3	1390	2	S51364	sperm tail-specifi
38	132.5	4.3	1070	2	T06733	kinesin homolog F2
39	131	4.2	1076	2	JC2217	major surface glyco
40	130	4.2	1175	2	PC3815	myosin heavy chain
41	129	4.2	1017	2	D34035	cell-cycle-depende
42	128.5	4.1	1051	2	JC4091	glycoprotein A - p
43	128	4.1	1005	2	A64465	hypothetical prote
44	128	4.1	3225	2	S52300	giantin - human
45	126.5	4.1	1927	2	A59236	embryonic muscle m
46	126.5	4.1	1972	1	A41604	myosin heavy chain
47	126.5	4.1	1974	2	T30010	hypothetical prote
48	126	4.1	1189	2	A54817	ATPase SC11, chrom
49	126	4.1	3259	1	A56539	giantin - human
50	124.5	4.0	1790	2	S67593	transport protein
51	124	4.0	2663	1	S28261	centromere protein
52	123.5	4.0	621	2	S10450	myosin heavy chain
53	123	4.0	1008	2	T30544	major surface glyco
54	123	4.0	1295	2	T24587	hypothetical prote
55	123	4.0	1538	2	T29095	cardiac muscle fac
56	123	4.0	1940	1	A24922	myosin heavy chain
57	122.5	3.9	344	2	S34153	mst101-1 protein -
58	122.5	3.9	1938	2	JCS421	smooth muscle myos
59	122.5	3.9	1972	2	JCS420	smooth muscle myos
60	122	3.9	779	2	T05990	hypothetical prote
61	122	3.9	2385	2	A34491	myosin heavy chain
62	122	3.9	2411	2	B34491	myosin heavy chain
63	121.5	3.9	932	2	S65214	probable alpha/gam
64	121.5	3.9	1280	2	A39117	170K lectin prec
65	121.5	3.9	1658	2	T42642	phosphoinositide 3
66	121	3.9	2253	2	T30336	nuclear/mitotic ap
67	120	3.9	1597	2	S68420	citrone - mouse
68	119.5	3.9	946	2	S28061	SCP1 protein - rat
69	119.5	3.9	1056	1	G02157	kinesin-like spind
70	119	3.8	1002	2	T30546	major surface glyco
71	119	3.8	1175	2	C35815	myosin heavy chain
72	119	3.8	1201	2	A35815	myosin heavy chain
73	119	3.8	2297	2	AB2494	hypothetical prote
74	118.5	3.8	1010	2	AH2553	hypothetical prote
75	118.5	3.8	1128	2	G86266	hypothetical prote
76	118.5	3.8	1899	2	T32732	PAM C-terminal int
77	118.5	3.8	1919	2	T42098	PAM interacting pr
78	118	3.8	700	2	S67610	probable membrane
79	118	3.8	1201	2	B35815	myosin heavy chain
80	117.5	3.8	963	1	A41919	kinesin heavy chain
81	117	3.8	911	2	S51441	hypothetical prote
82	117	3.8	993	2	S49461	synaptonemal compl
83	116.5	3.8	734	2	T27055	hypothetical prote
84	116	3.7	1282	2	JE0120	glycoprotein A - m
85	116	3.7	1313	2	F96673	hypothetical prote
86	116	3.7	1937	2	T38055	myosin heavy chain
87	116	3.7	2007	1	B43402	myosin heavy chain
88	116	3.7	2168	2	T30171	ninein - mouse
89	115.5	3.7	955	4	C40045	probable transcrip
90	115.5	3.7	1388	2	T30335	KLP2 protein - Afr
91	115.5	3.7	1875	2	S38173	myosin-like protei
92	115	3.7	576	1	B35128	DNA repair and gen
93	115	3.7	1061	2	C88690	protein F41H10.4 [
94	115	3.7	1642	2	T08880	NMDA receptor-bind
95	115	3.7	1727	2	T50073	myosin-like coiled
96	115	3.7	1961	1	A61231	myosin heavy chain
97	115	3.7	1976	2	A52532	myosin heavy chain
98	114.5	3.7	765	2	F96558	hypothetical prote
99	114.5	3.7	1136	2	F96564	hypothetical prote
100	114.5	3.7	1940	2	A29320	myosin heavy chain

# ALIGNMENTS

## RESULT 1

ABHUS  
serum albumin precursor [validated] - human  
N/Alternate names: preproalbumin  
N/Contains: kinetensin  
C/Species: Homo sapiens (man)  
C/Date: 29-Jul-1981 #sequence\_revision 31-Jan-1997 #text change 09-Jul-2004  
C/Accession: A93743; EMBL:U22768; I59286; I59313; G01747; S55314; A91420; S06422; S36  
R/Lawn, R.M.; Adelman, J.; Bock, S.C.; Franke, A.E.; Houck, C.M.; Najarian, R.C.; Seeburg  
Nucleic Acids Res. 9, 6103-6114, 1981  
A/Title: The sequence of human serum albumin cDNA and its expression in *Escherichia coli*  
A/Reference number: A93743; MUID:82081882; PMID:6171778  
A/Accession: A93743  
A/Molecule type: mRNA  
A/Residues: 1-419, 'K', 421-609 <LAW>  
A/Cross-references: UNIPROT:P02768; EMBL:V00495; GB:J00078; GB:L00133; NID:G2  
R/Dugaiczkyk, A.; Law, S.W.; Dennison, O.E.  
Proc. Natl. Acad. Sci. U.S.A. 79, 71-75, 1982  
A/Title: Nucleotide sequence and the encoded amino acids of human serum albumin mRNA.  
A/Reference number: A93936; MUID:82105994; PMID:6275391  
A/Accession: A93936  
A/Molecule type: mRNA  
A/Residues: 1-120, 'G', 122-609 <BUG>  
A/Cross-references: EMBL:V00494; NID:G28589; PIDN:CAA23753.1; PID:G28590  
R/Urano, Y.; Watanabe, K.; Sakai, M.; Tamaoki, T.  
J. Biol. Chem. 261, 3244-3251, 1986  
A/Title: The human albumin gene. Characterization of the 5' and 3' flanking regions and  
A/Reference number: I39427; MUID:86140099; PMID:2419329  
A/Accession: I39427  
A/Status: translation not shown  
A/Molecule type: DNA  
A/Residues: 1-26 <URA>  
A/Cross-references: GB:M13075; NID:G178330; PIDN:AAA51688.1; PID:G553173  
R/Watkins, S.; Madison, J.; Galliano, M.; Minchiotti, L.; Putnam, F.W.  
Proc. Natl. Acad. Sci. U.S.A. 91, 2275-2279, 1994  
A/Title: A nucleotide insertion and frameshift cause analbuminemia in an Italian family.  
A/Reference number: I59286; MUID:94181575; PMID:8134387  
A/Accession: I59286  
A/Status: translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 282-290, 'KSRFDLQ' <WAT>  
A/Cross-references: GB:S69192; NID:G546032; PIDN:AAB30282.1; PID:G546033  
R/Madison, J.; Galliano, M.; Watkins, S.; Minchiotti, L.; Porta, F.; Rossi, A.; Putnam,  
Proc. Natl. Acad. Sci. U.S.A. 91, 6476-6480, 1994  
A/Title: Genetic variants of human serum albumin in Italy: point mutants and a carboxyl-  
A/Reference number: I59313; MUID:94294404; PMID:8022807  
A/Accession: I59313  
A/Status: translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 589-590, 'ALPRRVKNLLQVKLP' <MAD>  
A/Cross-references: GB:S70799; NID:G547231; PIDN:AAB31177.1; PID:G547232  
A/Note: this frame-shift variant is designated albumin Bazzano; four additional variants  
R/Menaya, J.; Parrilla, R.; Ayuso, M.S.  
submitted to the EMBL Data Library, March 1995  
A/Reference number: G08292  
A/Accession: G08292  
A/Status: translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-120, 'G', 122-455 <VEN>  
A/Cross-references: EMBL:U22961; NID:G763428; PIDN:AAA64922.1; PID:G763431  
R/Ledgerwood, E.C.; George, P.M.; Peach, R.J.; Brennan, S.O.  
Biochem. J. 308, 321-325, 1995  
A/Title: Endoproteolytic processing of recombinant proalbumin variants by the yeast *Kex2*  
A/Reference number: S55314; MUID:9575251; PMID:7755581  
A/Accession: S55314  
A/Molecule type: protein  
A/Residues: 19-27 <LED>  
R/Meloun, B.; Moravsek, L.; Kostka, V.  
FEBS Lett. 58, 134-137, 1975

A/Title: Complete amino acid sequence of human serum albumin.  
A/Reference number: A91420; MUID:76187907; PMID:1225573  
A/Accession: A91420  
A/Molecule type: protein  
A/Residues: 25-117, 'EQ', 120-154, 'Q', 156-193, 'E', 195-387, 'H', 389-390, 'Y', 392-393, 'A', 395-  
R/Roehr, U.; Spitteller, G.; Tripier, D.  
Justus Liebig's Ann. Chem. 9, 881-884, 1988  
A/Title: Isolation and structure elucidation of middle-molecular weight peptides from ur  
A/Reference number: S06422  
A/Note: this paper is in German, with an English abstract  
A/Accession: S06422  
A/Molecule type: protein  
A/Residues: 25-48 <ROS>  
R/Finch, J.W.; Crouch, R.K.; Knapp, D.R.; Schey, K.L.  
Arch. Biochem. Biophys. 305, 595-599, 1993  
A/Title: Mass spectrometric identification of modifications to human serum albumin treat  
A/Reference number: S36882; MUID:93384321; PMID:8373198  
A/Accession: S36882  
A/Molecule type: protein  
A/Residues: 45-67, 141-160, 311-337, 469-490, 570-581 <PIN>  
R/Kausler, E.; Spitteller, G.  
Biol. Chem. Hoppe-Seyler 372, 849-855, 1991  
A/Title: Bruchstuecke aus Albumin und beta(2)-Mikroglobulin - Bestandteile der Mittelmol  
A/Reference number: S17599; MUID:92126241; PMID:1772598  
A/Accession: S17599  
A/Molecule type: protein  
A/Residues: 25-54, 354-357, 431-447 <KAU>  
A/Note: 49-Leu was also found  
R/Carraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.  
J. Immunol. 143, 1680-1684, 1989  
A/Title: Structures of histamine-releasing peptides formed by the action of acid proteas  
A/Reference number: A45800; MUID:89341406; PMID:2474609  
A/Accession: A45800  
A/Molecule type: protein  
A/Residues: 166-173, 'L' <CAR>  
R/Mogard, M.H.; Kobayashi, R.; Chen, C.F.; Lee, T.D.; Reeve Jr., J.R.; Shively, J.E.; Wa  
Biochem. Biophys. Res. Commun. 136, 983-988, 1986  
A/Title: The amino acid sequence of kinetensin, a novel peptide isolated from pepsin-tre  
A/Reference number: A03239; MUID:86242180; PMID:3087352  
A/Accession: A03239  
A/Molecule type: protein  
A/Residues: 166-173, 'L' <MOG>  
R/Galliano, M.; Minchiotti, L.; Porta, F.; Rossi, A.; Ferri, G.; Madison, J.; Watkins, S.  
Proc. Natl. Acad. Sci. U.S.A. 87, 8721-8725, 1990  
A/Title: Mutations in genetic variants of human serum albumin found in Italy.  
A/Reference number: A38255; MUID:91062352; PMID:2247440  
A/Accession: A38255  
A/Molecule type: protein  
A/Residues: 76-111 <GAL1>  
A/Accession: B38255  
A/Molecule type: protein  
A/Residues: 82-105, 'K', 107-110 <GAL2>  
A/Note: this variant is designated albumin Vibo Valentia  
A/Accession: A38255  
A/Molecule type: protein  
A/Residues: 76-83, 'K', 85-106 <GAL3>  
A/Note: this variant is designated albumin Torino  
R/Minchiotti, L.; Galliano, M.; Zapponi, M.C.; Fenni, R.  
Eur. J. Biochem. 214, 437-444, 1993  
A/Title: The structural characterization and bilirubin-binding properties of albumin Her  
A/Reference number: S33298; MUID:93292504; PMID:8513793  
A/Accession: S33298  
A/Molecule type: protein  
A/Residues: 255-263, 'E', 265-281 <MIN1>  
A/Note: this variant is designated albumin Herborn  
R/Minchiotti, L.; Galliano, M.; Stoppini, M.; Ferri, G.; Crespeau, H.; Rochu, D.; Porta,  
Biochim. Biophys. Acta 1119, 232-238, 1992  
A/Title: Two albumins with identical electrophoretic mobility are produced by differ  
A/Reference number: S21078; MUID:92190239; PMID:1347703  
A/Accession: S21078  
A/Molecule type: protein  
A/Residues: 354-356, 'K', 358-378 <MIN2>  
A/Note: this variant is designated albumin Sondrio; another variant Paris-2 is reported,

R:He, X.M.; Carter, D.C.  
 Nature 358, 209-215, 1992  
 A:Title: Atomic structure and chemistry of human serum albumin.  
 A:Reference number: A46756; MUID:92334427; PMID:1630489  
 A:Contents: annotation; X-ray crystallography, 2.8 angstroms  
 R:Brown, J.R.; Shockley, P.; Behrens, P.Q.  
 in The Chemistry and Physiology of the Human Plasma Proteins, Bing, D.H., ed., pp.23-40,  
 A:Reference number: A94442  
 A:Contents: annotation; three-dimensional structure and disulfide bonds  
 R:Saber, M.A.; Stockbauer, P.; Moravsek, L.; Meloun, B.  
 Collect. Czech. Chem. Commun. 42, 564-579, 1977  
 A:Title: Disulfide bonds in human serum albumin.  
 A:Reference number: A90930  
 A:Contents: annotation; disulfide bonds  
 R:Jacobsen, C.  
 Biochem. J. 171, 453-459, 1978  
 A:Title: Lysine residue 240 of human serum albumin is involved in high-affinity binding  
 A:Reference number: A90299; MUID:78186630; PMID:656055  
 A:Contents: annotation; bilirubin-binding site  
 R:Peters, T.; Reed, R.G.  
 in Albumin: Structure, Biosynthesis, Function, Peters, J., and Sjöholm, I., eds., 11-20,  
 A:Title: Serum albumin: conformation and active sites.  
 A:Reference number: A94408  
 A:Contents: annotation; binding sites  
 R:Harper, M.E.; Dugaiczky, A.  
 Am. J. Hum. Genet. 35, 565-572, 1983  
 A:Title: Linkage of the evolutionarily-related serum albumin and alpha-fetoprotein genes  
 A:Reference number: A90028; MUID:83279982; PMID:6192711  
 A:Contents: annotation; gene position  
 R:Walker, J.E.  
 FEBS Lett. 66, 173-175, 1976  
 A:Title: Lysine residue 199 of human serum albumin is modified by acetylsalicylic acid.  
 A:Reference number: A46755; MUID:76257808; PMID:955075  
 A:Contents: annotation  
 A:Note: the nonenzymatic transfer of an acetyl group from aspirin (acetylsalicylic acid  
 R:Bohney, J.P.; Fonda, M.L.; Felchhoff, R.C.  
 FEBS Lett. 298, 266-268, 1992  
 A:Title: Identification of Lys(190) as the primary binding site for pyridoxal 5'-phosphat  
 A:Reference number: A56294; MUID:92183881; PMID:1544460  
 A:Contents: annotation  
 A:Note: the nonenzymatic binding of pyridoxal phosphate to lysine-214 is described; in p  
 atase activity  
 C:Comment: Serum albumin, a predominant protein in the plasma of adults, is synthesized  
 in liver, protoporphyrin, long-chain fatty acids, prostaglandins, steroid hormones (weak  
 C:Comment: A large number of variants of human serum albumin have been described.  
 C:Genetics:  
 A:Gene: GDB:ALB  
 A:Cross-references: GDB:118990; OMIM:103600  
 A:Map position: 4q11-4q13  
 C:Superfamily: serum albumin; serum albumin repeat homology  
 C:Keywords: carrier protein; duplication; metal binding; phosphoprotein; plasma; pyridox  
 F:1-18/Domain: signal sequence #status predicted <SIG>  
 F:19-24/Domain: propeptide #status experimental <PRO>  
 F:25-609/Product: serum albumin #status experimental <MPT>  
 F:29-202/Domain: serum albumin repeat homology <SAI>  
 F:166-174/Product: kinetinsin #status experimental <KIP>  
 F:221-394/Domain: serum albumin repeat homology <SA2>  
 F:413-592/Domain: serum albumin repeat homology <SA3>  
 F:27/Binding site: copper (His) #status predicted  
 F:77-86,99-115,114-125,148-193,192-201,224-270,269-277,289-303,302-313,340-385,384-393,4  
 F:214/Binding site: pyridoxal phosphate (Lys) (covalent) #status experimental

Query Match 100.0%; Score 3103; DB 1; Length 609;  
 Best Local Similarity 100.0%; Pred. No. 9,3e-199;  
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 DVMCTAFHDNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180  
 DB 145 DVMCTAFHDNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 204  
 QY 181 KLDELURDSEKASSAKORLKCSLQKFGRAFKAWARELSORFPKAEFAEVSKLVTDLTK 240  
 DB 205 KLDELURDSEKASSAKORLKCSLQKFGRAFKAWARELSORFPKAEFAEVSKLVTDLTK 264  
 QY 241 VHTTECHGDLLECADRADLAKYICENQDSISSKLKECEKPLEKSHCHCIAEVENDEMPA 300  
 DB 265 VHTTECHGDLLECADRADLAKYICENQDSISSKLKECEKPLEKSHCHCIAEVENDEMPA 324  
 QY 301 DLPSLAADFVESKOVCKNYAEAKOVFLGMFLYEHARRHPDYVSVLLRLRAKYETTTLEKC 360  
 DB 325 DLPSLAADFVESKOVCKNYAEAKOVFLGMFLYEHARRHPDYVSVLLRLRAKYETTTLEKC 384  
 QY 361 CAAADPHCEYAKVDFEFKPLVEEPQNLKQNCLEFEQLGEYKFNALLVRYTKVPQVST 420  
 DB 385 CAAADPHCEYAKVDFEFKPLVEEPQNLKQNCLEFEQLGEYKFNALLVRYTKVPQVST 444  
 QY 421 PTLVEVSRLNGVSKCKGHPKAEKMPCAEDYLSVVLNQLCVLHKTPTVSDRVTKCCTES 480  
 DB 445 PTLVEVSRLNGVSKCKGHPKAEKMPCAEDYLSVVLNQLCVLHKTPTVSDRVTKCCTES 504  
 QY 481 LVNRRPCFSALEVDITYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVGHKPKAT 540  
 DB 505 LVNRRPCFSALEVDITYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVGHKPKAT 564  
 QY 541 KEQLKAVMDDFAAFVEKCKKADDDKTCFAEEGKLVAAASQAALGL 585  
 DB 565 KEQLKAVMDDFAAFVEKCKKADDDKTCFAEEGKLVAAASQAALGL 609

RESULT 2  
 A47391  
 Serum albumin precursor - rhesus macaque  
 C:Species: Macaca mulatta (rhesus macaque)  
 C:Date: 21-Jan-1994 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
 R:Watkins, S.; Sakamoto, Y.; Madison, J.; Davis, E.; Smith, D.G.; Dwulet, J.; Putnam, F  
 Proc. Natl. Acad. Sci. U.S.A. 90, 2409-2413, 1993  
 A:Title: cDNA and protein sequence of polymorphic macaque albumins that differ in bilir  
 A:Reference number: A47391; MUID:93211971; PMID:8460152  
 A:Contents: B/B homozygote  
 A:Accession: A47391  
 A:Status: preliminary  
 A:Molecule type: mRNA; protein  
 A:Residues: 1-600 <WAT>  
 A:Cross-references: UNIPROT:Q28522; GB:M90463; NID:G342294; PIDN:AAA36906.1; PID:G34229  
 A:Experimental source: liver  
 A:Note: sequence extracted from NCBI backbone (NCBI:128280, NCBIP:128281)  
 C:Superfamily: serum albumin; serum albumin repeat homology  
 F:21-194/Domain: serum albumin repeat homology <SA1>  
 F:213-386/Domain: serum albumin repeat homology <SA2>  
 F:405-584/Domain: serum albumin repeat homology <SA3>

Query Match 94.8%; Score 2942; DB 2; Length 600;  
 Best Local Similarity 93.5%; Pred. No. 4.6e-188;  
 Matches 545; Conservative 23; Mismatches 15; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFKDLGEENFKALVLIATAFYLOQCPEDHVKLVNEVTEFAKTCVADESAB 60  
 DB 17 DTHKEVAHRFKDLGEENFKALVLIATAFYLOQCPEDHVKLVNEVTEFAKTCVADESAB 76  
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQPERNECFLOHKDDNPILPRLVRPEV 120  
 DB 77 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQPERNECFLOHKDDNPILPRLVRPEV 136  
 QY 121 DVMCTAFHDNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180  
 DB 137 DVMCTAFHDNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 196

```
Qy 181 KLDELDEGKASSAKORLKCASLQKFGRAFKAWARELRSORFPAEAEVSKLVTDLTK 240
Dy 197 KLDELRDGSGKASSAKORLKCASLQKFGDRAFKAWARLRSQKFPKAEFAEYSKLVTDLTK 256
Qy 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPLLEKSHCIAEVENDEMPA 300
Dy 257 VHTCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPLLEKSHCIAEVENDEMPA 316
Qy 301 DLPSLAADFVSKDKVCKNYAEAKDVFLGMFLYEVARRHPDYSVLLRLAKTYETTLTKC 360
Dy 317 DLPSLAADYVSKDKVCKNYAEAKDVFLGMFLYEVARRHPDYSVLLRLAKYAEATLEKC 376
Qy 361 CAADAPHECYAKVFEDEFKPLVEEPONLIKQNCLEFEQLGEYKFQNALVRYTKVPQVST 420
Dy 377 CAADAPHECYAKVFEDEFKPLVEEPONLVKQNCLEFEQLGEYKFQNALVRYTKVPQVST 436
Qy 421 PTLVEVSRLGKVGSKCKCKHPEAKRMPCAEDYLSVLLNQLCVLHEKTPVSDRVTKCCTES 480
Dy 437 PTLVEVSRLGKVGAKCKCKLPEAKRMPCAEDYLSVLLNRLCVLHEKTPVSEKVTKCCTES 496
Qy 481 LVNRRPCFSALVEDEVTVVPKFNATETTFHADICTLSEKERQIKKOTALVELVXHKPKAT 540
Dy 497 LVNRRPCFSALVELDRAVYVPKFNATETTFHADICTLSEKERQIKKOTALVELVXHKPKAT 556
Qy 541 KEQLKAVMDDFAAAFVEKCKCKADDKETCFABEGKKLVAASQAAL 583
Dy 557 KEQLKGVMDNFAAAFVEKCKCKADDKETCFABEGPKFVAASQAAL 599
```

## RESULT 3

```
S57632
serum albumin precursor - cat
C:Species: Felis silvestris catus (domestic cat)
C>Date: 19-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C:Accession: J04660; S57632
R:Hiiger, C.; Grigioni, F.; Hentges, F.
Gene 169, 295-296, 1996
A:Title: Sequence of the gene encoding cat (Felis domesticus) serum albumin.
A:Reference number: J04660; MUID:96194624; PMID:8647469
A:Accession: J04660
A:Molecule type: mRNA
A:Residues: 1-608 <H12>
A:Cross-references: UNIPROT:P49064; EMBL:X84842; NID:G886484; PIDN:CAA59279.1; PID:G8864
A:Experimental source: liver
C:Comment: This protein is the major protein component in plasma. It functions as a mult
ein has 35 conserved cysteine residues.
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: liver; plasma
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-24/Domain: propeptide #status predicted <PRP>
F:25-608/Product: serum albumin #status predicted <MAT>
F:29-202/Domain: serum albumin repeat homology <SA1>
F:221-394/Domain: serum albumin repeat homology <SA2>
F:413-592/Domain: serum albumin repeat homology <SA3>
```

```
Query Match 84.4%; Score 2620; DB 2; Length 608;
Best Local Similarity 82.0%; Pred. No. 1.2e-166;
Matches 478; Conservative 52; Mismatches 53; Indels 0; Gaps 0;
```

```
Qy 1 DAHKSEVAHRPKDGEENFKALVLIAPAYLQQCFPFEDHVKLVNEVTEFAKTCVADESAAE 60
Dy 25 EAHQSEIAHRENDLGEEHFRLVLVAFSQYLQQCFPFEDHVKLVNEVTEFAKTCVADESAA 84
Qy 61 NCDKSLHTLFGDKLCTVATLRETGYEMADCCAKOEPERNECFLOHKDNNPLRLVRREV 120
Dy 85 NCEKSLHLLGDKLCTVATLRETGYEMADCCAKOEPERNECFLOHKDNNPLRFGQLVTPEA 144
Qy 121 DVMCTAFHNDNETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLTP 180
Dy 145 DAMCTAFHNEQRFLGKYLVEIARRHPYFYAPPELLFYAEYKGVFTCECEAADKAACLTP 204
Qy 181 KLDELDEGKASSAKORLKCASLQKFGRAFKAWARELRSORFPAEAEVSKLVTDLTK 240
Dy 197 KLDELRDGSGKASSAKORLKCASLQKFGDRAFKAWARLRSQKFPKAEFAEYSKLVTDLTK 256
```

```
Db 205 KVDALREKVLASSAKERLKCASLQKFGRAFKAWARELRSORFPAEAEISKLVTDLAK 264
Qy 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPLLEKSHCIAEVENDEMPA 300
Dy 265 IHKECCCHGDLLECCADRADLAKYICENQDSISTKLKECCGKPVLEKSHCISEVERDELUPA 324
Qy 301 DLPSLAADFVSKDKVCKNYAEAKDVFLGMFLYEVARRHPDYSVLLRLAKTYETTLTKC 360
Dy 325 DLPLPLAVDFVEDKEVKYQEAQDVLGTFLYEVSRHPEYSVLLRLAKEYEATLEKC 384
Qy 361 CAADAPHECYAKVFEDEFKPLVEEPONLIKQNCLEFEQLGEYKFQNALVRYTKVPQVST 420
Dy 385 CATDDPPACYAHVFEDEFKPLVEEPNLVKTNCELEFEKLGEYGFQNALVRYTKVPQVST 444
Qy 421 PTLVEVSRLGKVGSKCKCKHPEAKRMPCAEDYLSVLLNQLCVLHEKTPVSDRVTKCCTES 480
Dy 445 PTLVEVSRLGKVGSKCKCTHPEABRLSCAEDYLSVLLNRLCVLHEKTPVSRVTKCCTES 504
Qy 481 LVNRRPCFSALVEDEVTVVPKFNATETTFHADICTLSEKERQIKKOTALVELVXHKPKAT 540
Dy 505 LVNRRPCFSALQVDEVTVVPKFSATETTFHADICTLPEAEKQIKKQSALVELLXHKPKAT 564
Qy 541 KEQLKAVMDDFAAAFVEKCKCKADDKETCFABEGKKLVAASQAAL 583
Dy 565 BEQLKTMGDFGFSFVDCCKAAEDKEACFAEBEGPKLVAASQAAL 607
```

## RESULT 4

## ABHOS

```
serum albumin precursor - horse
C:Species: Equus caballus (domestic horse)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C:Accession: S34053
R:Ho, J.X.; Holowachuk, E.W.; Norton, E.J.; Twigg, P.D.; Carter, D.C.
Eur. J. Biochem. 215, 205-212, 1993
A:Title: X-ray and primary structure of horse serum albumin (Equus caballus) at 0.27-nm
A:Reference number: S34053; MUID:93345495; PMID:8344282
A:Accession: S34053
A:Molecule type: mRNA
A:Residues: 1-607 <HOA>
A:Cross-references: UNIPROT:P35747; GB:X74045; NID:G399671; PIDN:CAA52194.1; PID:G399672
C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,
teroid hormones (weak bonds with these hormones promote their transfer across the membra
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: carrier protein; duplication; metal binding; plasma
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-24/Domain: propeptide #status predicted <PRO>
F:25-607/Product: serum albumin #status predicted <MAT>
F:29-201/Domain: serum albumin repeat homology <SA1>
F:220-393/Domain: serum albumin repeat homology <SA2>
F:412-591/Domain: serum albumin repeat homology <SA3>
F:27/Binding site: copper (His) #status predicted
F:77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,4
F:263/Binding site: bilirubin (Iys) #status predicted
```

```
Query Match 79.8%; Score 2475.5; DB 1; Length 607;
Best Local Similarity 76.3%; Pred. No. 4.8e-157;
Matches 445; Conservative 69; Mismatches 68; Indels 1; Gaps 1;
```

```
Qy 1 DAHKSEVAHRPKDGEENFKALVLIAPAYLQQCFPFEDHVKLVNEVTEFAKTCVADESAAE 60
Dy 25 DTHKSEIAHRENDLGEKHFKGLVLVAFSQYLQQCFPFEDHVKLVNEVTEFAKTCVADESAAE 84
Qy 61 NCDKSLHTLFGDKLCTVATLRETGYEMADCCAKOEPERNECFLOHKDNNPLRLVRREV 120
Dy 85 NCDKSLHTLFGDKLCTVATLRETGYEMADCCAKOEPERNECFLOHKDNNPLRKL-KEEP 143
Qy 121 DVMCTAFHNDNETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLTP 180
Dy 144 DAQCAAQEDDPDKFLGKYLVEIARRHPYFYGPPELLFHAEEYKADFTCECPADDKLAACLIP 203
Qy 181 KLDELDEGKASSAKORLKCASLQKFGRAFKAWARELRSORFPAEAEVSKLVTDLTK 240
Dy 197 KLDELRDGSGKASSAKORLKCASLQKFGDRAFKAWARLRSQKFPKAEFAEYSKLVTDLTK 256
```

Db 204 KLDALKERILLSSAKERLKCSSFQNGPGERAVKAVSARLSQKPKADFAEVSKIYTDLTk 263  
QY 241 VTECHGDLLECADRADLAKYICENQDSISSKLKECEKPELLEKSHCIAEVENDEMPA 300  
Db 264 VHKECHGDLLECADRADLAKYICENQDSISSKLKECEKPELLEKSHCIAEVENDLPS 323  
QY 301 DPLSLAADPVESKDVCKYAEAKDVLFGHFLYFYARRHDPYVSVLLRLAKTYETLEKC 360  
Db 324 DLPALAADPAEDKCKYKADKDPVFLGFLYFYARRHDPYVSVLLRLAKTYETLEKC 383  
QY 361 CAAADPHCYAKVDFDEKFLVPEPONLIIKONCEPEQGEYKFQNALLVRYTKVPQVST 420  
Db 384 CEAADPPACYRIVFDQFTPLVEEPKSLVKNCNDLPEEVEGYDFQNALIVRYTKKAPQVST 443  
QY 421 PTLVEVRNLGVKGVKCKKHPKAPKMPCAEDYLSVVLNOLCVLHCKTPVSDRVTKCCTES 480  
Db 444 PTLVEIGRTVLGVKGRCKKLPESERLPCSENLHALLNRLCVLHCKTPVSEKITKCTDS 503  
QY 481 LVNRPRCSALEVDETYVPEKFAETFTFHADICTLSEKERQIKKQATLVELVKKPKAT 540  
Db 504 LAERPPCSALBDEGVVPKPEKATFTTFHADICTLPEDEKQIKKQALAEVLVGHKPKAT 563  
QY 541 KEQLKAVMDDFAAFEVKECKKADDKETCFPAEEGKLVAAQOAL 583  
Db 564 KEQLKTVLGNFSAFAVKACCGREDKACFAEEGPKLVASSQLAL 606

RESULT 5  
ABBS  
serum albumin precursor [validated] - bovine  
N/Alternate names: 67K protein; preproalbumin  
C/Species: Bos primigenius taurus (cattle)  
C/Date: 24-Apr-1984 #sequence revision 30-Sep-1993 #text change 09-Jul-2004  
C/Accession: A38885; A36401; A91258; B60808; S10780; D45800; A26693; A90309; A91458; A94551  
R/Holowachuk, E.W.; Stoltenborg, J.K.; Reed, R.G.; Peters Jr., T.  
Submitted to the EMBL Data Library, August 1991  
A/Description: Bovine serum albumin: cDNA sequence and expression.  
A/Reference number: A38885  
A/Accession: A38885  
A/Molecule type: mRNA  
A/Residues: 1-607 <HOL>  
A/Cross-references: UNIPROT:P04277; EMBL:M73215  
R/Hitayama, K.; Akashi, S.; Furuya, M.; Fukuhara, K.  
Biochem. Biophys. Res. Commun. 173, 639-646, 1990  
A/Title: Rapid confirmation and revision of the primary structure of bovine serum albumin  
A/Reference number: A36401; MUID:91083649; PMID:2260975  
A/Accession: A36401  
A/Molecule type: protein  
A/Residues: 25-41, 'H', 43-189, 'E', 191-213, 'T', 215-323, 'D', 325-393, 'TS', 396-607 <HIR>  
R/MacGillivray, R.T.A.; Chung, D.W.; Davie, E.W.  
Eur. J. Biochem. 98, 477-485, 1979  
A/Title: Biosynthesis of bovine plasma proteins in a cell-free system.  
A/Reference number: A91258; MUID:80024278; PMID:488109  
A/Accession: A91258  
A/Molecule type: protein  
A/Residues: 1-32 <MAG>  
R/Hsieh, J.C.; Lin, F.P.; Tam, M.F.  
Anal. Biochem. 170, 1-8, 1988  
A/Title: Electrophoretic onto glass-fiber filter from an analytical isoelectrofocusing gel  
A/Reference number: A60808; MUID:88267456; PMID:3389500  
A/Accession: B60808  
A/Molecule type: protein  
A/Residues: 25-41 <HSI>  
R/Strawich, E.; Glimcher, M.J.  
Eur. J. Biochem. 191, 47-56, 1990  
A/Title: Tooth 'enamelins' identified mainly as serum proteins. Major 'enamelin' is albumin  
A/Reference number: S10780; MUID:90336641; PMID:2379503  
A/Accession: S10780  
A/Molecule type: protein  
A/Residues: 25-41, 'H', 43-57, 59-64 <STR>  
R/Carraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.  
J. Immunol. 143, 1680-1684, 1989  
A/Title: Structures of histamine-releasing peptides formed by the action of acid proteases

A/Reference number: A45800; MUID:89341406; PMID:2474609  
A/Accession: D45800  
A/Molecule type: protein  
A/Residues: 163-172 <CAR>  
R/Carraway, R.E.; Mitra, S.P.; Cochran, D.E.  
J. Biol. Chem. 262, 5968-5973, 1987  
A/Title: Structure of a biologically active neurotensin-related peptide obtained from pm  
A/Reference number: A26693; MUID:87194805; PMID:2437111  
A/Accession: A26693  
A/Molecule type: protein  
A/Residues: 165-172, 'L', 'CA2'  
R/Reed, R.G.; Putnam, F.W.; Peters Jr., T.  
Biochem. J. 191, 867-868, 1980  
A/Title: Sequence of residues 400-403 of bovine serum albumin.  
A/Reference number: A90309; MUID:82023364; PMID:7283978  
A/Accession: A90309  
A/Molecule type: protein  
A/Residues: 402-433 <REE>  
R/Brown, J.R.  
Fed. Proc. 34, 591, 1975  
A/Title: Structure of bovine serum albumin.  
A/Reference number: A91458  
A/Accession: A91458  
A/Molecule type: protein  
A/Residues: 25-41, 'H', 43-117, 'EQ', 120-179, 181-189, 'E', 191-194, 'A', 196-213, 'T', 215-288, 'P'  
R/Brown, J.R.  
submitted to the Atlas, April 1975  
A/Reference number: A94551  
A/Accession: A94551  
A/Molecule type: protein  
A/Residues: 190-195 <BR2>  
R/Brown, J.R.  
Fed. Proc. 33, 1389, 1974  
A/Reference number: A91457  
A/Contents: annotation: disulfide bonds  
R/Werlen, R.C.; Offord, R.E.; Rose, K.  
Biochem. J. 302, 907-911, 1994  
A/Title: Preparation and characterization of novel substrates of insulin proteinase (EC  
A/Reference number: S55232; MUID:95031935; PMID:7945219  
A/Accession: S55232  
A/Status: preliminary  
A/Molecule type: protein  
A/Residues: 529-536; 569-572 <WER>  
C/Superfamily: serum albumin; serum albumin repeat homology  
C/Keywords: carrier protein; copper binding; duplication; plasma  
F.1-18/Domain: signal sequence #status experimental <SIG>  
F.19-24/Domain: propeptide #status experimental <PRO>  
F.25-607/Product: serum albumin #status experimental <MPT>  
F.29-201/Domain: serum albumin repeat homology <SA1>  
F.220-393/Domain: serum albumin repeat homology <SA2>  
F.412-591/Domain: serum albumin repeat homology <SA3>  
F.27/Binding site: copper (His) #status predicted  
F.77-86, 99-115, 114-125, 147-192, 191-200, 223-269, 268-276, 288-302, 301-312, 339-384, 383-392,

Query Match 78.8%; Score 2446.5; DB 1; Length 607;  
Best Local Similarity 75.6%; Pred. No. 4.1e-155;  
Matches 441; Conservative 71; Mismatches 70; Indels 1; Gaps 1;

QY 1 DAHKEVAHRPKDLGEENFKALVLTAFAYQLQCCPFEDHVKLVNNEVTEFAKTCVADESAE 60  
Db 25 DTHKSEIAHRPDKLGESEQFKGLVLIAFSQYLQCCPFEDHVKLVNNEVTEFAKTCVADES 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPNERNECFLOHKDNNPLRLVRP 120  
Db 85 GCEKSLHTLFGDELCKVLSLRETYGDMADCCCKQEPNERNECFSLSHKDDSPDLK- KPDP 143  
QY 121 DMCTAFHDNRETFLKKVLYEIAIRHPYFYAPELLFFAKRYKAAATECCQAADKACLLP 180  
Db 144 NTLCDKFADEKKEKFWGKLYEIAIRHPYFYAPELLFYANKYNGVFDQCCQAEKDGACLLP 203  
QY 181 KLDLREDSGKASSAKORLKCKASLOKFGERAFKAWAVARLSORFFPKAEAEVSKLVTDLTK 240  
Db 204 KIETMRKVLASSAKQRURCASIQKFGERALKAWVARLSQKPKAEFEVTKLVTDLTK 263









[illegible]

RESULT 9  
 JC5838  
 albumin - Mongolian jird  
 C/Species: Meriones unguiculatus (Mongolian jird)  
 C/Date: 05-Mar-1998 #sequence\_revision 13-Mar-1998 #text\_change 09-Jul-2004  
 C/Accession: JC5838  
 R/Yoshida, K.; Sato-Ohshima, A.; Sinohara, H.  
 DNA Res. 4, 351-354, 1997  
 A/Title: Sequencing of cDNA encoding serum albumin and its extrahepatic synthesis in the  
 A/Reference number: JC5838; MUID:98116663; PMID:9455485  
 A/Accession: JC5838  
 A/Molecule type: mRNA  
 A/Residues: 1-609 <YOS>  
 A/Cross-references: UNIPROT:O35090; DBJ:AB006197; NID:G2317277; PIDN:BA21765.1; PID:92  
 A/Experimental source: liver  
 C/Superfamily: serum albumin; serum albumin repeat homology  
 F:222-395/Domain: serum albumin repeat homology <SA2>

RESULT 10  
A05139  
serum albumin - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 05-Jun-1987 #sequence revision 17-Mar-2000 #text change 09-Jul-2004

C;Accession: A05139; I48638  
R;Minghetti, P.P.; Law, S.W.; Dugaiczak, A.  
Mol. Biol. Evol. 2, 347-358, 1985  
A;Title: The rate of molecular evolution of alpha-fetoprotein approaches that of pseudogenes  
A;Reference number: A93055; MUID:88216123; PMID:2452956  
A;Accession: A05139  
A;Molecule type: mRNA  
A;Residues: 1-418 <MIN>  
A;Cross-references: UNIPROT:P07724; GB:M16111; NID:G9191764; PIDN:AAA37190.1; PID:G9191765  
R;Boccaccio, C.; Deschattrette, J.; Meunier-Rotival, M.  
Gene 88, 181-186, 1990  
A;Title: Empty and occupied insertion site of the truncated LINE-1 repeat located in the 3' noncoding region of the human alpha-fetoprotein gene  
A;Reference number: I48638; MUID:90269606; PMID:1971802  
A;Accession: I48638  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 379-453 <BC>  
A;Cross-references: EMBL:X13060; NID:G52939; PIDN:CAA311458.1; PID:G8899334  
C;Superfamily: serum albumin; serum albumin repeat homology  
C;Keywords: carrier protein; duplication; metal binding; plasma  
F;1-104/Domain: serum albumin repeat homology (fragment) <SA1>  
F;123-296/Domain: serum albumin repeat homology <SA2>  
F;315-453/Domain: serum albumin repeat homology (fragment) <SA3>

```

RESULT 11
ABCS
serum albumin precursor - chicken
C/Species: Gallus gallus (chicken)
C/Date: 31-Dec-1993 #sequence revision 31-Dec-1993 #text_change 09-Jul-2004
C/Accession: S15571; A05078; A13451
R/Cassidy, A. I.; Salkild, C. K.; Baverstock, P.; Wallace, J. C.
submitted to the EMBL Data Library, July 1991
A/Reference number: S15571
A/Accession: S15571
A/Molecule type: mRNA
A/Residues: 1-615 <CAS>

```





```
QY 417 QVSTPTLVEVSRNLGKVGSKCKHPKMPKCAEDYLSVVLNOLCVLHKTPTVSDRVTKC 476
Db 441 QLTSSLMATRKMAATAATCCQLSEDKLLACGEAAADIIIGHLCIRHEMTVPNPGVQC 500
QY 477 CTESLVNRRPCFSALEVDEYVYKFEFNAETFTFHADICTLSEKEROIKKQTALVELVKHK 536
Db 501 CTSSYANRRPCFSLVVDYVYPPAFSDDKFIHKDLCOAQGVALQTMKQEFNLINLVKQK 560
QY 537 PKATKEQLKAVMDDFAAFEVKCKKADDKETCFEAEGKLVVAASQAALGL 585
Db 561 PQITERQLEAVIADFSGLLEKCCQGEQEVCFEAEGQKLISKTRTALGV 609

RESULT 14
FGO
alpha-fetoprotein precursor - gorilla
C:Species: Gorilla gorilla (gorilla)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C:Accession: A37970
R:Ryan, S.C.; Zielinski, R.; Dugaiczkyk, A.
Genomics 9, 60-72, 1991
A:Title: Structure of the gorilla alpha-fetoprotein gene and the divergence of primates.
A:Reference number: A37970; MUID:91169517; PMID:1706310
A:Accession: A37970
A:Molecule type: DNA
A:Residues: 1-609 <RYA>
A:Cross-references: UNIPROT:P28050; GB:M38272; NID:g817963; PID:AAAY3520.1; PID:g177041
C:Genetics:
A:Map position: 4q11-12
A:Introns: 29/1; 46/2; 90/3; 205/3; 238/2; 281/3; 353/2; 397/3; 430/2; 476/3; 551/3
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: embryo; fetus; globulin; glycoprotein; metal binding; plasma
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-609/Product: alpha-fetoprotein #status predicted <SA1>
F:29-202/Domain: serum albumin repeat homology <SA1>
F:221-394/Domain: serum albumin repeat homology <SA2>
F:413-592/Domain: serum albumin repeat homology <SA3>
F:422/Binding site: copper (His) #status predicted
F:99-114,113-124,148-193,192-224-270,269-277,289-303,302-313,384-393,416-462,461-472
F:249/Binding site: bilirubin (Lys) #status predicted
F:251/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 40.0%; Score 1242.5; DB 1; Length 609;
Best Local Similarity 39.8%; Pred. No. 4.5e-75;
Matches 233; Conservative 117; Mismatches 232; Indels 7; Gaps 3;
QY 3 HKSE-----VAHRFKDLGGENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADE 57
Db 22 HRNEYGIALDSYQCTAISLADLATIIFAQVQBATYKYSKWKDALTAIEKPTGDE 81
QY 58 SAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPNERNECFLOHKDDNP-NLPLRV 116
Db 82 QSAGCLENPAPLEELCHEKEILEKYG-LSDCSSQSEGRHNCFLAHKPTPASIPFQ 140
QY 117 REVDVMTAFHDNETFLKYLVEIARHPYAPPELLFFAKRYKAAATECCQADKAA 176
Db 141 VPEVPTSCAYEEDRETFMNKFYEIARHPFLYAPTILLAAARYDKIIPSCCKAENAYE 200
QY 177 CLLPKDLDELDEKAGSQRKLCASLOKFGERAFKAWAVARLSQRFPKAEFVYSKLYT 236
Db 201 CFQTKAATYKELRESLLNQHCACVMKQFGRTFOAITVKLSQKFTKNTVTEIOKLVL 260
QY 237 DLTKVHTECHGDLLEACADRDALAKYICENQDISSKLKECECKPFLLEKSHCIAREVD 296
Db 261 DVAHVHEHCGRGDLVLDQGEKIMSYICSQQDTLSNKKITECCKLTTLERGQCIIHAEND 320
QY 297 EMPADILPSLAADPVESKDVCKNYAEAKDVFLGNFLYEYARRHPDYSVLLLRLLAKYET 356
Db 321 EKEGSLPNLNRFLGRDQNFQSSGSEKNIPLASFVHEYSRRHPQLAVSVILRVAKGYQSL 380
QY 357 LEKCAAAADPHCEYAKVDFEFKPLVEEPQNLIKONCELFQGEYKFNQALLVRYTKVP 416
Db 381 LEKCFOTENPLECDKGEBELQKYIQESQALAKSCGLFQKLEYYLQNAFLVAYTKCAP 440
```

```
QY 417 QVSTPTLVEVSRNLGKVGSKCKHPKMPKCAEDYLSVVLNOLCVLHKTPTVSDRVTKC 476
Db 441 QLTSSLMATRKMAATAATCCQLSEDKLLACGEAAADIIIGHLCIRHEMTVPNPGVQC 500
QY 477 CTESLVNRRPCFSALEVDEYVYKFEFNAETFTFHADICTLSEKEROIKKQTALVELVKHK 536
Db 501 CTSSYANRRPCFSLVVDYVYPPAFSDDKFIHKDLCOAQGVALQTMKQEFNLINLVKQK 560
QY 537 PKATKEQLKAVMDDFAAFEVKCKKADDKETCFEAEGKLVVAASQAALGL 585
Db 561 PQITERQLEAVIADFSGLLEKCCQGEQEVCFEAEGQKLISKTRTALGV 609

RESULT 15
ABXL72
74K albumin precursor - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C:Accession: B41682; S02693; A05288
R:Moskaitis, J.E.; Sargent, T.D.; Smith Jr., L.H.; Pastori, R.L.; Schoenberg, D.R.
Mol. Endocrinol. 3, 464-473, 1989
A:Title: Xenopus laevis serum albumin: sequence of the complementary deoxyribonucleic a-
e during development.
A:Reference number: A41682; MUID:89313788; PMID:2747653
A:Accession: B41682
A:Molecule type: mRNA
A:Residues: 3-607 <MOS>
A:Cross-references: UNIPROT:P14872; GB:M21442; NID:g213930; PID:AAA49637.1; PID:g21393
R:Schorpp, M.; Doebebling, U.; Wagner, U.; Ryffel, G.U.
J. Mol. Biol. 199, 83-93, 1988
A:Title: 5'-flanking and 5'-proximal exon regions of the two Xenopus albumin genes. Del
A:Reference number: S02692; MUID:88172470; PMID:2451026
A:Accession: S02693
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-48 <SCH>
A:Cross-references: EMBL:226826
R:Wolfe, A.P.; Glover, J.F.; Martin, S.C.; Tenniswood, M.P.R.; Williams, J.L.; Tata, J.
Eur. J. Biochem. 146, 489-496, 1985
A:Title: Deinduction of transcription of Xenopus 74-kDa albumin genes and destabilizati
A:Reference number: A05288; MUID:85126974; PMID:3971963
A:Accession: A05288
A:Molecule type: mRNA
A:Residues: 459-502, 'L', 504-557 <WOL>
A:Cross-references: GB:M28276
A:Note: the authors translated the codon TAT for residue 63 as Thr
C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,
mones (weak bonds with these hormones promote their transfer across the membranes), thy
C:Genetics:
A:Introns: 27/1
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: carrier protein; duplication; glycoprotein; metal binding; plasma
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-24/Domain: propeptide #status predicted <PRO>
F:25-607/Product: 74K serum albumin #status predicted <SA1>
F:320-393/Domain: serum albumin repeat homology <SA2>
F:412-591/Domain: serum albumin repeat homology <SA3>
F:430/Binding site: copper (His) #status predicted
F:80-88,101-117,116-127,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392
F:256/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 38.8%; Score 1205; DB 1; Length 607;
Best Local Similarity 39.3%; Pred. No. 1.4e-72;
Matches 227; Conservative 108; Mismatches 239; Indels 4; Gaps 2;
QY 3 HKSEVAHRFKDLGGENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADEAENC 62
Db 30 HHKHIADVYVTALETRTFKGLTLAIVSQNLKQCSLELSKLVNEINDFAKSCINDKTPE-C 88
QY 63 DKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPNERNECFLOHKDDNP-NLPLRVREVDV 122
```

```
Db 89 EKPVGTLFFDKLCADPAVGAVNYEWSKECAKQDPERAQCFKAHRDHEHT---SIKPEEE 145
Qy 123 MCTAFHDNEETFLKYLVEIARRHPYFYAPELLFFAKYKAAFTTECCQAADKAACLPLKL 182
Db 146 TCKLLKEHPDLLSAFIEEARNHPDLPPYPAVLATKQYKHLAEHCCEEDKEKCFSEKM 205
Qy 183 DELRDEGKASSAKQRLKCSAQKQGERAFKAWAVARLSQRPFPKAFBAEVSKLVDLTGVH 242
Db 206 KQLMKQSHSIEDKQHHFCWILNDNPFKVLKALNARVSHRPKAFKLAHNFTVEVTHFI 265
Qy 243 TECCHGDLLECADRADLAKYICENQDSISSKLECCCKPLEKSHCIAEVENDEMPADL 302
Db 266 KDCCHDDMFECMTERLELTHECTQHKBSLKKLECCNIPLLERTYICIVLTLENDVPAEL 325
Qy 303 PSLAADFVESKDVCKNYAEKDVFLGMFLYEVARRHPDYSVVLRLAKTYETTLKCCA 362
Db 326 SOPITEFTEDPHVCYSKAENNEVFGRYLHAVSRKHQELSBOFLQSAKYESLNNCKCK 385
Qy 363 AADPHECYAKVDFEFPKLVPEEPQNLIKQNCLEFQOLGEYKFQNALLVRYTKVPQVSTPT 422
Db 386 TDNPFECYKGNDRPMNAEKERFAVLKQNCILHEHGYLFENELLIRYTKMPQVSDET 445
Qy 423 LVEVSRNLGKVGSKCKHPEAKRMPCAEYLSVLNQLCVLHEKTPVSDRVTKCTESLV 482
Db 446 LIGIAHQWADIGEHCACVAPENQRMPCAEGLDTILIGKMCERQKKTFINNHVHACCTDSYS 505
Qy 483 NRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKQATVLELVKHKPKATKE 542
Db 506 GMRSCFTALGPDDEDYVPPVTDTHFDKDKICTANDKEKHQIKQKFLKLIKVSQKLEKN 565
Qy 543 OLKAWMDDFAPFVEKCKKADKETCFABEKGKLVAAQS 580
Db 566 HIDECSAEFLKXWQKCTADEHQPCDFTEKPVLIHQ 603

RESULT 16
I46986
A:Species: Canis lupus familiaris (dog)
C:Date: 04-Sep-1997 #sequence_revision 07-Nov-1997 #text_change 20-Aug-1999
C:Accession: I46986
R:Spitzauer, S.; Schweiger, C.; Sperr, W.R.; Pandjaitan, B.; Valent, P.; Muhl, S.; Ebner
J. Allergy Clin. Immunol. 93, 614-627, 1994
A:Title: Molecular characterization of dog albumin as a cross-reactive allergen.
A:Reference number: I46986; PMID:94201492; PMID:7512102
A:Accession: I46986
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-265 <SPI>
A:Cross-references: GB:S72946; NID:g633937; PIDN:AAB30434.1; PID:g633938
C:Superfamily: serum albumin; serum albumin repeat homology
F:7-180/Domain: serum albumin repeat homology <SA2>

Query Match 38.1%; Score 1181.5; DB 2; Length 265;
Best Local Similarity 82.6%; Pred. No. 1.9e-71;
Matches 218; Conservative 23; Mismatches 22; Indels 1; Gaps 1;
Qy 192 SSAKORLKCSAQKQGERAFKAWAVARLSQRPFPKAFBAEVSKLVDLTGVHTECHGDL 251
Db 2 SSAKERFKCSAQKQGERAFKAWAVARLSQRPFPKADFAEISKVVDLTGVHKECHGDL 61
Qy 252 ECADRADLAKYICENQDSISSKLECCCKPLEKSHCIAEVENDEMPADLPSLAADFVE 311
Db 62 ECADRADLAKYICENQDSISITLKECCCKPVLEKSQLAEVERDELPGDLPSLAADFVE 121
Qy 312 SKDVCKNYAEKDVFLGMFLYEVARRHPDYSVVLRLAKTYETTLKCCAADPHECYA 371
Db 122 KDEVCKNYQEAQKDVFLGFLYEVSRHPEYSVSLRLAKEYEATLEKCCATDDPPPTCYA 181
Qy 372 KVFDEFPKLVPEEPQNLIKQNCLEFQOLGEYKFQNALLVRYTKVPQVSTPTL-VEVSRNL 430
Db 182 KYLDEFPKLVDPBQNLVNTCNELFEKLGEGYQNALLVRYTKAPQVSTPTLVVEVSRKL 241
```

```
Qy 431 GKVGSKCKHPEAKRMPCAEYLS 454
Db 242 GKVGTKCKRPESERMSCADDPLS 265

RESULT 17
ABXL68
A:Species: Xenopus laevis (African clawed frog)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C:Accession: A41682; S02692
R:Moskatis, J.E.; Sargent, T.D.; Smith Jr., L.H.; Pastori, R.L.; Schoenberg, D.R.
Mol. Endocrinol. 3, 464-473, 1989
A:Title: Xenopus laevis serum albumin: sequence of the complementary deoxyribonucleic ac
e during development.
A:Reference number: A41682; MUID:89313788; PMID:2747653
A:Accession: A41682
A:Molecule type: mRNA
A:Residues: 1-608 <MOS>
A:Cross-references: UNIPROT:P08759; GB:M18350
R:Schorpp, M.; Doebebling, U.; Wagner, U.; Ryffel, G.U.
J. Mol. Biol. 199, 83-93, 1988
A:Title: 5'-flanking and 5'-proximal exon regions of the two Xenopus albumin genes. Dele
A:Reference number: S02692; MUID:88172470; PMID:2451026
A:Accession: S02692
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-48 <SCH>
A:Cross-references: EMBL:Z26825
C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,
mones (weak bonds with these hormones promote their transfer across the membranes), thyr
C:Genetics:
A:Introns: 27/1
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: carrier protein; duplication; metal binding; plasma
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-24/Domain: propeptide #status predicted <PRO>
F:25-608/Product: 68K serum albumin #status predicted <MAT>
F:32-203/Domain: serum albumin repeat homology <SA1>
F:221-394/Domain: serum albumin repeat homology <SA2>
F:413-592/Domain: serum albumin repeat homology <SA3>
F:30/Binding site: copper (His) #status predicted
F:80-89,102-118,117-128,148-193,192-201,224-270,269-277,289-303,302-313,340-385,384-393,

Query Match 37.9%; Score 1175.5; DB 1; Length 608;
Best Local Similarity 38.3%; Pred. No. 1.3e-70;
Matches 222; Conservative 111; Mismatches 240; Indels 7; Gaps 2;
Qy 3 HKSEVAHRFKDGLGFENFKALVLIAPQYLOQCPEDHVKLVNEVTEFAKTCVADESANC 62
Db 30 HHKHIADMYNLLTERTFKGLTLAIVSQNLQKCSLEELSKLVNEINDFAKSGTGNDKTPEC 89
Qy 63 DKSLLHTLFGDKLCTVATLRETYGEMADCCAKQBERNECFLOHK--DDNPNLPLRVREP 120
Db 90 EKPITGLFYDKLCAADPKVGVNYSWKSCSKQDPERAQCFRAHRVFEHNP----VRPKP 144
Qy 121 DVMCTAHDEETFLKYLVEIARRHPYFYAPELLFFAKRYKAAFTTECCQAADKAACL 180
Db 145 EETCALFKHPDDLLSAFIEEARNHPDLPPYPAVLTLTQQYKLVVEHCCEEDKDKCPAE 204
Qy 181 KLDELIRDEGKASSAKQRLKCSAQKQGERAFKAWAVARLSQRPFPKAFBAEVSKLVDLT 240
Db 205 KMKELMKHSHSIEDKQHHFCWILNDNPFKVLKALNARVSHRPKAFKLAHKTTEETH 264
Qy 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLECCCKPLEKSHCIAEVENDEMPA 300
Db 265 FIKDCCHGDMFECMTERLELTHECTQHKDELSTLKECCNIPLLERTYICIVLTLENDVPA 324
Qy 301 DLPSLAADFVESKDVCKNYAEKDVFLGMFLYEVARRHPDYSVVLRLAKTYETTLKCC 360
Db 325 ELSRPTTEFTEDPHVCYSKAENKESFLERSIPWQSQTPELSEQFLQSAKEYESLNNKC 384
Qy 361 CAAADPHECYAKVDFEFPKLVPEEPQNLIKQNCLEFQOLGEYKFQNALLVRYTKVPQVST 420
```





101 CHETELSNKYG - FSGCCNGSGVERHOCGLARKKTAADPSVPPFFHPETAESCPEAYENRAM 159

134 FLKKYLYEYARRHPFYAPPELLFFAKRYKAAFTTECCQAADKAAACLLPLKLDLDEDEGKASS 193

160 SINTFYIDVYSKRNPFLYAPITILVLAQYDKAVPACCKADNWESCFCOTKRASMAKELREGS 219

194 AKORLKLASLOKFGERAFAKAWAVARISQRPKPAFAEVSKLVLTDLTKVHTTECHGDLLLEC 253

220 MLNEHVCVIRKFGSRNLQAVLLIKLSQKPPKANITEIRKLALDVAHIHEQCCHNAMEC 279

254 ADDRADLAKYICENODSISSKLKECKECPKLPLEKSHCIAEVENDEMPADLPSLAADFVESK 313

280 LODGESVMTMHCQQBILSKTAECCKLPFTIEUGYCIHHAENGDKPEGLTLPNSEFLGDR 339

314 DVCKNVAEKDVFELGMFLVEYARRHPDYSVVLRLRLAKTETTTLEKCAAADPHECVAKV 373

340 NFAQFSSEKLLPMAEFLHEYSRHNPLPVSIVILKTAKSVOETLEKCSQSETSKCODNM 399

374 FDFEKLPEEPQNLIKONCELFEOQGEYKQNALLVRYTKKVPQVSTPTLVVEYSRNLGKV 433

400 EEBELQKHIOESQALAKOSCNLYKQLGPFYVQLNLFILGYTRKAPQLTSAELIDLTKGMVSI 459

434 GSKCKCHPEAKRNPCEADYLSVVLNOLCVLHEKTPVSDRVTKCTESLVNRRPCFSALEV 493

460 ASTCCQLSEKRSACGEGLADIIYHGLLRHEANPVNSGINHCSSSYSNRRLCITISFLR 519

494 DETYVPKEFNAETFTPHADICTLSEKERQIKKOTALVELVKHPKATKEQLKAVMDDFAA 553

520 DETYVPPPSSEDKFIHKDLCOAQGRALQTMKQELLINLVKQKPEMTEEOHAAVTADFSG 579

554 FVBCKCKADDKETCTFAEEGKKLVAAASQAALGL 585

580 LLEKCKKDDQOEACFAKEGPKLISKTRREALGV 611

RESULT 20

A54906

A:afamin precursor - human

N:Alternate names: alpha-albumin

C:Species: Homo sapiens (man)

C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999, #text\_change 09-Jul-2004

C:Accession: A54906; JG6143; S68554; S78082; I39425

R:Lichenstein, H.S.; Lyons, D.E.; Wurfel, M.M.; Johnson, D.A.; McGinley, M.D.; J. Biol. Chem. 269, 18149-18154, 1994

A:Title: Afamin is a new member of the albumin, alpha-fetoprotein, and vitamin A:Reference number: A54906; MUID:94239534; PMID:7517938

A:Accession: A54906

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-599 <Lit>

A:Cross-references: UNIPROT:P43652; GB:L32140; NID:G533885; PIDN:AAA21612.1; PI

R:Mishio, H.; Dugaiczky, A. Proc. Natl. Acad. Sci. U.S.A. 93, 7557-7561, 1996

A:Title: Complete structure of the human alpha-albumin gene, a new member of the A:Reference number: JG6143; MUID:96353855; PMID:8755513

A:Accession: JG6143

A:Molecule type: DNA

A:Residues: 1-599 <MS>

A:Cross-references: GB:U51243; NID:gl418261; PIDN:AAC50720.1; PID:gl418262

R:Nishio, H.; Heiskanen, M.; Palotie, A.; Belanger, L.; Dugaiczky, A. J. Mol. Biol. 259, 113-119, 1996

A:Title: Tandem arrangement of the human serum albumin multigene family in the A:Reference number: S68554; MUID:96240683; PMID:8648639

A:Accession: S68554

A:Molecule type: DNA

A:Residues: 1-29 <Niw>

A:Cross-references: GB:U51243; NID:gl418261

A>Note: neither the complete nucleic acid sequence nor the complete translation

R:Nishio, H.; Heiskanen, M.; Palotie, A.; Belanger, L.; Dugaiczky, A. submitted to the EMBL Data Library, March 1996

A:Description: Tandem arrangement of the human serum albumin multigene family i A:Reference number: S78082

A:Accession: S78082





**THIS PAGE BLANK (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2005, 08:39:21 ; Search time 178 Seconds  
(without alignments)  
1682.957 Million cell updates/sec

Title: US-10-816-042-18  
Perfect score: 3103  
Sequence: 1 DAHKEVAHRFDLGEENFK.....TCFAEEGKKLVAAASQAALGL 585

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : UniProt 03: \*  
1: uniprot\_sprot: \*  
2: uniprot\_trembl: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	3103	100.0	609	1	ALBU HUMAN
2	3079	99.2	609	2	Q68DN5
3	2942	94.8	600	1	ALBU MACMU
4	2620	84.4	608	1	ALBU FELCA
5	2614	84.2	584	2	Q7YSG3
6	2562	82.6	608	1	ALBU CANFA
7	2504	80.7	608	2	Q9SVB7
8	2475.5	79.8	607	1	ALBU HORSE
9	2457	79.2	608	1	ALBU RABIT
10	2450.5	79.0	607	1	ALBU BOVIN
11	2432.5	78.4	607	1	ALBU SHEEP
12	2426	78.2	608	1	ALBU RAT
13	2411.5	77.7	605	1	ALBU_PIG
14	2404.5	77.5	607	2	Q6GNH7
15	2387	76.9	608	2	Q6WDN9
16	2379.5	76.7	583	1	ALBU MERUN
17	2379.5	76.7	583	2	Q6B3Z0
18	2378	76.6	608	1	ALBU MOUSE
19	2374	76.5	608	2	Q8C7H3
20	2336	75.3	576	2	Q8C7C7
21	1991	64.2	417	2	Q8GYG0
22	1865.5	60.1	396	2	Q8IUK7
23	1557.5	50.2	615	1	ALBU CHICK
24	1295.5	41.7	527	2	Q8JIA9
25	1253.5	40.4	609	1	FETA_PANTR
26	1249.5	40.3	609	1	FETA_HUMAN
27	1242.5	40.0	609	1	FETA_GORGO
28	1242	40.0	609	2	Q8MJU5
29	1242	40.0	626	2	Q8UW05
30	1218.5	39.3	610	2	Q8MJ76
31	1205	38.8	607	1	ALB2_XENLA
					P14872 xenopus lae

32	1200	38.7	609	1	FETA_HORSE
33	1199	38.6	607	2	Q642P7
34	1183.5	38.1	608	2	Q7TSF3
35	1164.5	37.5	606	1	ALB1_XENLA
36	1087	35.0	624	2	Q8UW06
37	1084	34.9	605	1	FETA_MOUSE
38	1082	34.9	605	2	Q8BK65
39	1076	34.7	605	2	Q8BK56
40	1067	34.4	611	1	FETA_RAT
41	1055	34.0	599	1	AFAM_HUMAN
42	1045	33.7	400	2	Q8JIA7
43	955	30.8	603	2	Q9YGH6
44	944	30.4	611	1	AFAM_MOUSE
45	928.5	29.9	614	2	Q9I134
46	928	29.9	608	2	AFAM_RAT
47	888	28.6	406	2	Q8JIA8
48	771	24.8	167	2	Q6UXK4
49	747.5	24.1	608	1	ALB1_SALSA
50	742.5	23.9	608	1	ALB2_SALSA
51	739	23.8	205	2	Q8CG74
52	713	23.0	417	2	Q8ROJ9
53	699	22.5	382	1	ALBU_RANCA
54	440.5	14.2	1423	1	ALBU_PETMA
55	386	12.4	474	1	VTDB_HUMAN
56	385	12.4	476	2	Q68FY4
57	381	12.3	476	1	VTDB_RAT
58	378	12.2	476	1	VTDB_RABIT
59	376.5	12.1	484	2	Q9W6F5
60	372	12.0	476	1	VTDB_MOUSE
61	365	11.8	551	2	Q42279
62	331	10.7	122	2	Q9WZ8
63	290	9.3	123	2	Q9WZ6
64	284	9.2	464	2	Q6DGV8
65	203	6.5	1894	2	Q75XB6
66	200	6.4	1927	2	Q75WX9
67	198.5	6.4	1820	2	Q75XM5
68	196	6.3	807	2	Q83ZS7
69	195	6.3	1810	2	Q75X88
70	195	6.3	2002	2	Q75XH1
71	194.5	6.3	1813	2	Q75XE3
72	192.5	6.2	1797	2	Q75X61
73	190.5	6.1	1795	2	Q75XJ8
74	188	6.1	1723	2	Q75XM8
75	186.5	6.0	1796	2	Q75XQ2
76	184	5.9	1819	2	Q9ZLV0
77	184	5.9	1927	2	Q25262
78	183	5.9	1802	2	Q75X05
79	177.5	5.7	1768	2	Q75X33
80	162.5	5.2	44	2	Q95MC2
81	162.5	5.2	680	2	Q9V6S8
82	161	5.2	1079	2	Q96V11
83	156	5.0	1026	2	Q74669
84	154.5	5.0	3843	2	Q9USD0
85	153.5	4.9	661	2	Q8MS79
86	153.5	4.9	3843	2	Q9VU94
87	153	4.9	62	2	Q8M11
88	152.5	4.9	810	2	Q64JV4
89	151.5	4.9	1605	1	RRB1_MOUSE
90	150.5	4.9	1077	2	Q6AHS1
91	150	4.8	1065	2	Q01828
92	149	4.8	40	2	Q9TRA5
93	149	4.8	897	2	Q64JW5
94	148	4.8	804	2	Q62PS0
95	148	4.8	1028	2	Q74668
96	145	4.7	1825	2	Q7RIT2
97	144.5	4.7	8407	2	Q7ETM4
98	144.5	4.7	8797	1	SNE1_HUMAN
99	143	4.6	1069	2	Q6AHV5
100	142.5	4.6	218	2	Q66RQ0

ALIGNMENTS

RESULT 1  
ALBU HUMAN  
ID \_ALBU HUMAN STANDARD; PRT; 609 AA.  
AC P02768; O95574; Q13140; Q9P157; Q9P117; Q9UHS3; Q9UJ20;  
DT 21-JUL-1986 (Rel. 01, Created)  
DD 01-APR-1990 (Rel. 14, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Serum albumin precursor.  
GN Name=ALB;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86196112; PubMed=3009475;  
RA Minghetti P.P., Ruffner D.E., Kuang W.J., Dennison O.E., Hawkins J.W.,  
RT Beattie W.G., Dugaiczak A.;  
RT "Molecular structure of the human albumin gene is revealed by  
RT nucleotide sequence within g11-22 of chromosome 4.";  
RL J. Biol. Chem. 261:6747-6757(1986).  
RN [2]  
RP SEQUENCE FROM N.A., AND VARIANT LYS-420.  
RX MEDLINE=82081882; PubMed=6171778;  
RA Lawn R.M., Adelman J., Bock S.C., Franke A.E., Houck C.M.,  
RA Najarian R.C., Seeburg P.H., Wion K.L.;  
RT "The sequence of human serum albumin cDNA and its expression in E.  
RT coli.";  
RL Nucleic Acids Res. 9:6103-6114(1981).  
RN [3]  
RP SEQUENCE FROM N.A., AND VARIANT GLY-121.  
RX MEDLINE=82105994; PubMed=6275391;  
RA Dugaiczak A., Law S.W., Dennison O.E.;  
RT "Nucleotide sequence and the encoded amino acids of human serum  
RT albumin mRNA.";  
RL Proc. Natl. Acad. Sci. U.S.A. 79:71-75(1982).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RA Yang S., Zhang R.A., Qi Z.W., Yuan Z.Y.;  
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A. (PRO0903/PRO1708/PRO2044/PRO2619/PRO2675).  
RC TISSUE=Fetal liver;  
RA Zhang C., Yu Y., Zhang S., Wei H., Bi J., Zhou G., Dong C., Zai Y.,  
RA Xu W., Gao F., Liu M., He F., Zhang Y., Ouyang S., Luo L.;  
RT "Functional prediction of the coding sequences of 121 new genes  
RT deduced by analysis of cDNA clones from human fetal liver.";  
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A., AND VARIANT HIROSHIMA-1 LYS-378.  
RA Huang M.C., Wu H.T.;  
RT "The cDNA sequences of human serum albumin.";  
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
RN [7]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver, and Skeletal muscle;  
RX MEDLINE=23388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.J., Peters G.J., Abramson R.D., Mullighy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,  
RA Whitting M., Madan A.C., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krawinski M.I., Skalska U., Smallos D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [8]  
RP SEQUENCE OF 25-609.  
RX MEDLINE=7617907; PubMed=1225573; DOI=10.1016/0014-5793(75)80242-0;  
RA Meloun B., Moravsek L., Kostka V.;  
RT "Complete amino acid sequence of human serum albumin.";  
RL FEBS Lett. 58:134-137(1975).  
RN [9]  
RP SEQUENCE OF 25-609.  
RA Brown J.R., Shockley P., Behrens P.O.;  
RL (In) Bing D.H. (eds.);  
RL The chemistry and physiology of the human plasma proteins, pp.23-40,  
RL Pergamon Press, New York (1979).  
RN [10]  
RP SEQUENCE OF 1-455 FROM N.A.  
RC TISSUE=Liver;  
RA Menaya J., Parrilla R., Ayuso M.S.;  
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
RN [11]  
RP SEQUENCE OF 1-26 FROM N.A.  
RX MEDLINE=86140099; PubMed=2419329;  
RA Urano Y., Watanabe K., Sakai M., Tamaoki T.;  
RT "The human albumin gene. Characterization of the 5' and 3' flanking  
RT regions and the polymorphic gene transcripts.";  
RL J. Biol. Chem. 261:3244-3251(1986).  
RN [12]  
RP SEQUENCE OF 222-229.  
RX MEDLINE=76257808; PubMed=955075; DOI=10.1016/0014-5793(76)80496-6;  
RA Walker J.E.;  
RT "Lysine residue 199 of human serum albumin is modified by  
RT acetylsalicylic acid.";  
RL FEBS Lett. 66:173-175(1976).  
RN [13]  
RP SEQUENCE OF 25-44 AND 480-499.  
RC TISSUE=Heart;  
RX MEDLINE=95203287; PubMed=7895732;  
RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;  
RT "The human myocardial two-dimensional gel protein database: update  
RT 1994.";  
RL Electrophoresis 15:1459-1465(1994).  
RN [14]  
RP DISULFIDE BONDS.  
RA Saber M.A., Stockbauer P., Moravsek L., Meloun B.;  
RT "Disulfide bonds in human serum albumin.";  
RL Collect. Czech. Chem. Commun. 42:564-579(1977).  
RN [15]  
RP BILIRUBIN-BINDING SITE.  
RX MEDLINE=78186630; PubMed=656055;  
RA Jacobsen C.;  
RT "Lysine residue 240 of human serum albumin is involved in high-  
RT affinity binding of bilirubin.";  
RL Biochem. J. 171:453-459(1978).  
RN [16]  
RP VARIANT CANTERBURY ASN-337.  
RX MEDLINE=87157744; PubMed=3828358; DOI=10.1016/0167-4838(87)90088-4;  
RA Brennan S.O., Herbert P.;  
RT "Albumin Canterbury (313 Lys-->Asn). A point mutation in the second  
RT domain of serum albumin.";  
RL Biochim. Biophys. Acta 912:191-197(1987).  
RN [17]  
RP VARIANTS NAG-2 AND NAG-3.  
RX MEDLINE=88068523; PubMed=3479777;  
RA Takahashi N., Takahashi Y., Isobe T., Putnam F.W., Fujita M.,  
RA Satoh C., Neel J.V.;  
RT "Amino acid substitutions in inherited albumin variants from  
RT Amerindian and Japanese populations.";  
RL Proc. Natl. Acad. Sci. U.S.A. 84:8001-8005(1987).  
RN [18]

RP VARIANTS NAG-1; HIR-1; HIR-2 AND TOCHIGI.  
RX MEDLINE=89345611; PubMed=2762316;  
RA Arai K., Madison J., Huss K., Ishioka N., Satoh C., Fujita M.,  
RA Neel J.V., Sakurabayashi I., Putnam F.W.;  
RT "Point substitutions in Japanese alloalbumins."  
RL Proc. Natl. Acad. Sci. U.S.A. 86:6092-6096(1989).  
RN [19]  
RP VARIANTS MANAUS; OSAKA; NAGOYA; FUKUOKA; HONOLULU AND NEW-GUINEA.  
RX MEDLINE=90115905; PubMed=2404284;  
RA Arai K., Madison J., Shimizu A., Putnam F.W.;  
RA "Point substitutions in albumin genetic variants from Asia."  
RL Proc. Natl. Acad. Sci. U.S.A. 87:497-501(1990).  
RN [20]  
RP DESCRIPTION OF VARIANT REDHILL.  
RX MEDLINE=90115852; PubMed=2104980;  
RA Brennan S.O., Myles T., Peach R.J., Donaldson D., George P.M.;  
RT "Albumin Redhill (-1 Arg, 320 Ala-->Thr): a glycoprotein variant of  
RT human serum albumin whose precursor has an aberrant signal peptidase  
RT cleavage site."  
RL Proc. Natl. Acad. Sci. U.S.A. 87:26-30(1990).  
RN [21]  
RP VARIANTS TORINO LYS-84; VARESE HIS-23 AND VIBO VALENTIA LYS-106.  
RX MEDLINE=91062352; PubMed=2247440;  
RA Galliano M., Minchiotti L., Porta F., Rossi A., Ferri G., Madison J.,  
RA Watkins S., Putnam F.W.;  
RT "Mutations in genetic variants of human serum albumin found in  
RT Italy."  
RL Proc. Natl. Acad. Sci. U.S.A. 87:8721-8725(1990).  
RN [22]  
RP VARIANT VENEZIA.  
RX MEDLINE=91296740; PubMed=2068071;  
RA Watkins S., Madison J., Davis E., Sakamoto Y., Galliano M.,  
RA Minchiotti L., Putnam F.W.;  
RT "A donor splice mutation and a single-base deletion produce two  
RT carboxyl-terminal variants of human serum albumin."  
RL Proc. Natl. Acad. Sci. U.S.A. 88:5959-5963(1991).  
RN [23]  
RP VARIANTS IOWA CITY-2 VAL-25; IOWA CITY-1 VAL-389; KOMAGOME-3 HIS-23;  
RX KOMAGOME-2 ARG-152 AND KOMAGOME-1 GLU-396.  
RX MEDLINE=92052189; PubMed=1946412;  
RA Madison J., Arai K., Feld R.D., Kyle R.A., Watkins S., Davis E.,  
RA Matsuda Y.-I., Amaki I., Putnam F.W.;  
RT "Genetic variants of serum albumin in Americans and Japanese."  
RL Proc. Natl. Acad. Sci. U.S.A. 88:9853-9857(1991).  
RN [24]  
RP VARIANT CASEBROOK ASN-518.  
RX MEDLINE=91316157; PubMed=1859851; DOI=10.1016/0925-4439(91)90023-3;  
RA Peach R.J., Brennan S.O.;  
RT "Structural characterization of a glycoprotein variant of human serum  
RT albumin: albumin Casebrook (494 Asp-->Asn)."  
RL Biochim. Biophys. Acta 1097:49-54(1991).  
RN [25]  
RP VARIANTS SONDRIO LYS-357 AND PARIS-2 ASN-587.  
RX MEDLINE=92190239; PubMed=1347703; DOI=10.1016/0167-4838(92)90207-T;  
RA Minchiotti L., Galliano M., Stoppini M., Ferri G., Crespeau H.,  
RA Rochu D., Porta F.;  
RT "Two alloalbumins with identical electrophoretic mobility are produced

Query Match 100.0%; Score 3103; DB 1; Length 609;

Best Local Similarity 100.0%; Pred. No. 5.1e-188;

Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFKDLGEENFKALVLIATAFYLOQCPEDHVKLVNEVTEFAKTCVADESAE 60

DB 25 DAHKEVAHRFKDLGEENFKALVLIATAFYLOQCPEDHVKLVNEVTEFAKTCVADESAE 84

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDPNLPRLVRPEV 120

DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDPNLPRLVRPEV 144

QY 121 DVMCTAFHNEETFLKKLYEYIARRHPYFVAPPELLFFAKRYKAAFTCCCAADKAACLLP 180

DB 145 DVMCTAFHNEETFLKKLYEYIARRHPYFVAPPELLFFAKRYKAAFTCCCAADKAACLLP 204

QY 181 KUDELURDEGKASSAKQRLKASLOKFGGERAFKAWAVARLSQRPFPKAEPAEVSKLVTDLTk 240

DB 205 KUDELURDEGKASSAKQRLKASLOKFGGERAFKAWAVARLSQRPFPKAEPAEVSKLVTDLTk 264

QY 241 VHTECHGDLLECADRADLAKYICENODSISSKLKECCCKPPLLEKSHCIAEVENDEMPA 300

DB 265 VHTECHGDLLECADRADLAKYICENQDSISSKLKECCCKPPLLEKSHCIAEVENDEMPA 324

QY 301 DLPSLAADFVESKDVCKYKNAEAKDVFGLMFLYEYARRHPDYSVVLNLLRLAKTYETTTLEK 360

DB 325 DLPSLAADFVESKDVCKYKNAEAKDVFGLMFLYEYARRHPDYSVVLNLLRLAKTYETTTLEK 384

QY 361 CAAADPHECYAKVDFEFPPLVEEPQNLIKQNCLEFQELGEYKFNALLVRYTKKVPQVST 420

DB 385 CAAADPHECYAKVDFEFPPLVEEPQNLIKQNCLEFQELGEYKFNALLVRYTKKVPQVST 444

QY 421 PTLVEVSRNLGKVGSKCCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480

DB 445 PTLVEVSRNLGKVGSKCCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504

QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTPHADICTLSEKERQIKKQTLALVELVKKHKPKAT 540

DB 505 LVNRRPCFSALEVDETYVPKEFNAETFTPHADICTLSEKERQIKKQTLALVELVKKHKPKAT 564

QY 541 KEQLKAVMDDFAAAFVEKCKKADDDKTCFPAEKGKLVAAASQAALGL 585

DB 565 KEQLKAVMDDFAAAFVEKCKKADDDKTCFPAEKGKLVAAASQAALGL 609

# RESULT 2

Q68DN5

ID Q68DN5 PRELIMINARY; PRT; 609 AA.

AC Q68DN5;

DT 25-OCT-2004 (TremblRel. 28, Created)

DT 25-OCT-2004 (TremblRel. 28, Last sequence update)

DT 25-OCT-2004 (TremblRel. 28, Last annotation update)

DE Hypothetical protein DKFZp779N1935.

GN Name=DKFZp779N1935;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Liver;

RG The German cDNA Consortium;

RA Bloeker H., Boecker M., Brandt P., Mewes H.W., Weil B., Amid C.,

RA Osanger A., Fobo G., Han M., Wiemann S.;

RL Submitted (AUG-2004) to the EMBL/Genbank/DBJ databases.

DR EMBL; CR749331; CAH18185.1; -.

DR InterPro; IPR001703; AlphaFoldProt.

DR InterPro; IPR000264; Serum albumin.

DR Pfam; PF00273; Serum albumin; 3.

DR PRINTS; PR00803; AFETOPROTEIN.

DR PRINTS; PR00802; SERUMALBUMIN.

DR ProDom; PD002486; Serum albumin; 1.

DR SMART; SM00103; ALBUMIN; 3.

DR PROSITE; PS00212; ALBUMIN; 3.

KW Hypothetical protein.

SQ SEQUENCE 609 AA; 69402 MW; 3BA3AFF17BF99E94 CRC64;

Query Match 99.2%; Score 3079; DB 2; Length 609;

Best Local Similarity 99.1%; Pred. No. 1.7e-186;

Matches 580; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFKDLGEENFKALVLIATAFYLOQCPEDHVKLVNEVTEFAKTCVADESAE 60

DB 25 DAHKEVAHRFKDLGEENFKALVLIATAFYLOQCPEDHVKLVNEVTEFAKTCVADESAE 84

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDPNLPRLVRPEV 120

DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDPNLPRLVRPEV 144

```
QY 121 DVMCTAFHDNBEETFLKKYLYEIAIRRHYPFYAPPELLFFAKRYKAAAFTECCQAADKAACLIP 180
Db 145 DVMCTAFHDNBEETFLKKYLYEIAIRRHYPFYAPPELLFFAKRYKAAAFTECCQAADKAACLIP 204
QY 181 KLDELRLDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRFPAEFAEYVSKLVTDLT 240
Db 205 KLDELRLDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRFPAEFAEYVSKLVTDLT 264
QY 241 VHTCCHGDLLECADRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 300
Db 265 VHTCCHGDLLECADRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVSKQVCKNYAEAKOVFLGMFLYEYARRHPDYSVLLRLAKTYETTLK 360
Db 325 DLPSLAADFVSKQVCKNYAEAKOVFLGMFLYEYARRHPDYSVLLRLAKTYETTLK 384
QY 361 CAADPHCEYAKVDFEFPKLVBEQPNLIKQNCFLFEQLGEYKFNQALLVRYTKVPQVST 420
Db 385 CAADPHCEYAKVDFEFPKLVBEQPNLIKQNCFLFEQLGEYKFNQALLVRYTKVPQVST 444
QY 421 PTLVEVSNNLGKVGSKCCKHPKAPCAEDYLSVNLQNLVHLHKTTPVSRVTKCCTES 480
Db 445 PTLVEVSNNLGKVGSKCCKHPKAPCAEDYLSVNLQNLVHLHKTTPVSRVTKCCTES 504
QY 481 LVNRRPCFSALEVDVTPYVKEFNABETFTFHADICTLSEKERQIKKOTALVELVKHKPKAT 540
Db 505 LVNRRPCFSALEVDVTPYVKEFNABETFTFHADICTLSEKERQIKKQAAALVELVKHRPKAT 564
QY 541 KEQLKAVNDDFAAFVEKCKCKADKTCFAEBEGKLVAAASQAALGL 585
Db 565 KEQLKAVNDDFAAFVEKCKCKADKTCFAEBEGKLVAAASQAALGL 609
```

## RESULT 3

```
ALBU MACMU
ID ALBU MACMU STANDARD; PRT; 600 AA.
AC Q28522;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, last sequence update)
DT 05-JUL-2004 (Rel. 44, last annotation update)
DE Serum albumin precursor (Fragment).
GN Name=ALB;
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Macaca.
OX NCBI TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93211971; PubMed=8460152;
RA Watkins S.A., Sakamoto Y., Madison J.M., Davis E.M., Smith D.G.,
RA Dwulet J., Putnam F.W.;
RT "cDNA and protein sequence of polymorphic macaque albumins that differ
in bilirubin binding.";
RL proc. Natl. Acad. Sci. U.S.A. 90:2409-2413(1993).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
hormones, bilirubin and drugs. Its main function is the regulation
of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
```

```
DR 103 EMBL; M90463; AAA36906.1; -.
DR PIR; A47391; A47391.
DR HSSP; P02768; 1B7B.
DR InterPro; IPR001703; Alphafetoprot.
DR Pfam; PF00273; Serum albumin; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Copper; lipid-binding; Metal-binding; Repeat; Signal.
FT NON TER 1
FT SIGNAL <1 10 By similarity.
FT PROPEP 11 16 By similarity.
FT CHAIN 17 600 Serum albumin.
FT DOMAIN 17 197 Albumin 1.
FT DOMAIN 204 389 Albumin 2.
FT DOMAIN 396 587 Albumin 3.
FT METAL 19 19 Copper (By similarity).
FT BINDING 256 256 Bilirubin (potential).
FT DISULFID 69 78 By similarity.
FT DISULFID 91 107 By similarity.
FT DISULFID 106 117 By similarity.
FT DISULFID 140 185 By similarity.
FT DISULFID 184 193 By similarity.
FT DISULFID 216 262 By similarity.
FT DISULFID 261 269 By similarity.
FT DISULFID 281 295 By similarity.
FT DISULFID 294 305 By similarity.
FT DISULFID 332 377 By similarity.
FT DISULFID 376 385 By similarity.
FT DISULFID 408 454 By similarity.
FT DISULFID 453 464 By similarity.
FT DISULFID 477 493 By similarity.
FT DISULFID 492 503 By similarity.
FT DISULFID 530 575 By similarity.
FT DISULFID 574 583 By similarity.
SQ SEQUENCE 600 AA; 67880 MW; E45C871A670B740B CRC64;
```

Query Match 94.8%; Score 2942; DB 1; Length 600;

Best Local Similarity 93.5%; Pred. No. 7.5e-178;

Matches 545; Conservative 23; Mismatches 15; Indels 0; Gaps 0;

```
QY 1 DAHSEVAHRPKDLGEEFNKALVIAFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
Db 17 DTHSEVAHRPKDLGEEHFKGLVLAFSQYLQCCPFEEHVKLVNEVTEFAKTCVADESAAE 76
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPPLRLVRPEV 120
Db 77 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPPLRLVRPEV 136
QY 121 DVMCTAFHDNEETFLKKYLYEIAIRRHYPFYAPPELLFFAKRYKAAAFTECCQAADKAACLIP 180
Db 137 DVMCTAFHDNEETFLKKYLYEIAIRRHYPFYAPPELLFFAKRYKAAAFTECCQAADKAACLIP 196
QY 181 KLDELRLDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRFPAEFAEYVSKLVTDLT 240
Db 197 KLDELRLDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRFPAEFAEYVSKLVTDLT 256
QY 241 VHTCCHGDLLECADRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 300
Db 257 VHTCCHGDLLECADRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 316
QY 301 DLPSLAADFVSKQVCKNYAEAKOVFLGMFLYEYARRHPDYSVLLRLAKTYETTLK 360
Db 317 DLPSLAADFVSKQVCKNYAEAKOVFLGMFLYEYARRHPDYSVLLRLAKTYETTLK 376
QY 361 CAADPHCEYAKVDFEFPKLVBEQPNLIKQNCFLFEQLGEYKFNQALLVRYTKVPQVST 420
Db 377 CAADPHCEYAKVDFEFPKLVBEQPNLIKQNCFLFEQLGEYKFNQALLVRYTKVPQVST 436
QY 421 PTLVEVSNNLGKVGSKCCKHPKAPCAEDYLSVNLQNLVHLHKTTPVSRVTKCCTES 480
```

```

Db 437 PTLVSVNRNLGVAKCKCLPEAKRMPCAEDYLSVVLNRLCVLHKEKTPVSEKVTCCCTES 496
QY 481 LVNRRPCFSALEVEDYVPKEPNAETFTFHADICTLSEKERQIKKQTALVELVGHKPKAT 540
Db 497 LVNRRPCFSALEDEAYVPKAFNAETFTFHADMCTLSEKEKQVKQTALVELVGHKPKAT 556
QY 541 KEOLKAVMDDFAAFEKCKCKADDKETCFPAEKGKLVAAASQAAL 583
Db 557 KEOLKGVMDNFNAFVEKCKCKADDKACFAEKGKLVAAASQAAL 599

RESULT 4
ALBU_FELCA STANDARD; PRT; 608 AA.
AC P49064;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Serum albumin precursor (Allergen Fel d 2).
GN Name=ALB;
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96194824; PubMed=8647469; DOI=10.1016/0378-1119(95)00851-9;
RA Hilger C., Grigioni F., Kohnen M., Hentges F.;
RT "Sequence of the gene encoding cat (Felis domesticus) serum albumin.";
RL Gene 169:295-296(1996).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- ALLERGEN: Causes an allergic reaction in human.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; X84842; CAA59279.1; -.
CC DR PIR; J04660; S57632.
CC DR HSP; P02768; 1E7B.
CC
CC InterPro; IPR000264; Serum albumin.
CC DR Pfam; PF00273; Serum albumin; 3.
CC DR PRINTS; PR00802; SERUMALBUMIN.
CC DR ProDom; PD002486; Serum_albumin; 1.
CC DR SMART; SM00103; ALBUMIN; 3.
CC DR PROSITE; PS00212; ALBUMIN; 3.
CC
CC Allergen; Copper; Lipid-binding; Metal-binding; Repeat; Signal.
KW SIGNAL
FT 1 18 By similarity.
FT PROPEP 19 24 By similarity.
FT CHAIN 25 608 Serum albumin.
FT DOMAIN 25 205 Albumin 1.
FT DOMAIN 212 397 Albumin 2.
FT DOMAIN 404 595 Albumin 3.
FT METAL 27 27 Copper.
FT METAL 77 86 By similarity.
FT DISULFID 99 115 By similarity.
FT DISULFID 114 125 By similarity.
FT DISULFID 148 193 By similarity.
FT DISULFID 192 201 By similarity.
FT DISULFID 224 270 By similarity.
FT DISULFID 269 277 By similarity.
FT DISULFID 289 303 By similarity.

```

```

FT DISULFID 302 313 By similarity.
FT DISULFID 340 385 By similarity.
FT DISULFID 384 393 By similarity.
FT DISULFID 416 462 By similarity.
FT DISULFID 461 472 By similarity.
FT DISULFID 485 501 By similarity.
FT DISULFID 500 511 By similarity.
FT DISULFID 538 583 By similarity.
FT DISULFID 582 591 By similarity.
SQ SEQUENCE 608 AA; 68659 MW; 07B629CAC5F60E5F CRC64;

Query Match
Best Local Similarity 84.4%; Score 2620; DB 1; Length 608;
Matches 478; Conservative 52; Mismatches 53; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGSENFKALVIAFAQYLOQCQPFEDHVKLVNEVTEFAKTCVADESA 60
Db 25 EAHQSEIAHRFNDLGEHFRGLVAVFSQYLOQCQPFEDHVKLVNEVTEFAKGVADQSA 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCAKQBPNERNECFLOHKDDNPRLVRPEV 120
Db 85 NCEKSLHLLGDKLCTVATLRETYGEMADCAKQBPNERNECFLOHKDDNPFGVLTPEA 144
QY 121 DVMTAFHNDNEETFLKLYETARHPYFYAPPELLFFFAKRYKAAFTCCQAAAKAALLP 180
Db 145 DAMCTAFHENEORFLGKYLYETARHPYFYAPPELLFYAEYKGVFTECCAAADKAALLP 204
QY 181 KLDELURDGGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 240
Db 205 KVDALREKVLASSAKERLKCASLQKFGGERAFKAWAVARLSQRPKAEFAEVSKLVTDLAK 264
QY 241 VTECHGDLLECADRADLAKYICENODSISSKLKECEKPLLESKSHCIAEVENDEMPA 300
Db 265 IHKECHGDLLECADRADLAKYICENODSISSKLKECEKPLLESKSHCIAEVENDEMPA 324
QY 301 DLPSLAADPVSKDKVYAKQVFLGMLFVEYARRHPDYSVLLLRLLAKTYETLEK 360
Db 325 DLPPLAADPVSKDKVYAKQVFLGMLFVEYARRHPDYSVLLLRLLAKTYETLEK 384
QY 361 CAAADPHCYAKVDFEFKPLVEEPQNLKQNCLEFQELGEYKFNALLVRYTKVPQVST 420
Db 385 CATDPPACYAHVDFEFKPLVEEPQNLKQNCLEFQELGEYKFNALLVRYTKVPQVST 444
QY 421 PTLVSVNRNLGVAKCKCLPEAKRMPCAEDYLSVVLNRLCVLHKEKTPVSEKVTCCCTES 480
Db 445 PTLVSVNRNLGVAKCKCLPEAKRMPCAEDYLSVVLNRLCVLHKEKTPVSEKVTCCCTES 504
QY 481 LVNRRPCFSALEVEDYVPKEPNAETFTFHADICTLSEKERQIKKQTALVELVGHKPKAT 540
Db 505 LVNRRPCFSALEVEDYVPKEPNAETFTFHADICTLSEKERQIKKQTALVELVGHKPKAT 564
QY 541 KEOLKAVMDDFAAFEKCKCKADDKETCFPAEKGKLVAAASQAAL 583
Db 565 KEOLKAVMDDFAAFEKCKCKADDKETCFPAEKGKLVAAASQAAL 607

RESULT 5
QYSG3 PRELIMINARY; PRT; 584 AA.
ID QYSG3
AC QYSG3
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Albumin (Fragment).
GN Name=alb;
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Reininger R., Swoboda I., Bohle B., Hauswirth A.W., Valent P.,

```



RA Rumpold H., Valenta R., Spitzauer S.;  
RL Submitted (MAY-2002) to the EMBL/GenBank/DBSJ databases.

DR EMBL; AJ487677; CAD32275.1; -;  
DR HSSP; P02768; 1E7B.

DR GO; GO:0005615; C:extracellular space; IEA.

DR GO; GO:0005386; F:carrier activity; IEA.

DR GO; GO:0006810; P:transport; IEA.

DR InterPro; IPR000264; Serum albumin.

DR Pfam; PF00273; Serum\_albumin; 3.

DR ProDom; PD002486; Serum\_albumin; 1.

DR SMART; SM00103; ALBUMIN; 3.

DR PROSITE; PS00212; ALBUMIN; 3.

FT NON TER 1

SQ SEQUENCE 584 AA; 65908 MW; B51002F12902C9CE CRC64;

Query Match 84.2%; Score 2614; DB 2; Length 584;

Best Local Similarity 81.6%; Pred. No. 4e-157;

Matches 476; Conservative 54; Mismatches 53; Indels 0; Gaps 0;

QY 1 DAHSEVAHRRFDLGEENFKALVIAFAQYLQCCPFEDHVKLVNVEVTEFAKTCVADESAAE 60

DB 1 EAHGSEIAHRNDLGEERFGLVLFVAFSQYLQCCPFEDHVKLVNVEVTEFANGCVADQSA 60

QY 61 NCDKSLHTLFGDKLCTVATLRETTGEMADCCAKQEPERNECFLOHKDNDPNLPRLVPEV 120

DB 61 NCKSLHLELFGDKLCTVATLRETTGEMADCCAKQEPERNECFLOHKDNDPNPGFQLVTEA 120

QY 121 DVMCTAFHDNTEFLKVKLYIARIHPYFYAPELLFFAKRYKAAFTCCQAAADKAACLLP 180

DB 121 DAMCTAFHENEQRFKGLKLYIARIHPYFYAPELLFYAEEYRGVFTCECEAADKAACLTP 180

QY 191 KLDELDRSGKASSAKORLKASLOKFGGERAFKAVARLSORFPKABEAVSKLVTDLTG 240

DB 191 KYDALREKVLASSAKERLKASLOKFGGERAFKAVARLSORFPKADPAEISKLVTDLAK 240

QY 241 VHTCECHGDLLECADRADLAKYICENQDSISKLKECCERPLEKSKHICIAEVENDEMPA 300

DB 241 IHKECCHGDLLECADRADLAKYICENQDSISKLKECCGKPVLEKSKHICIEVERDELPA 300

QY 301 DLPLSADAFVSKDKVCKYAKADVFLGMFLYEYARRHPDYSVLLLLAKYIETTLK 360

DB 301 DLPLAADFVEDKEVKCKYQAKDVLFTGLFLEYSRHRPEYSVLSLLLAKEYEATLEK 360

QY 361 CAADPHCYAKVDFEPKPLVEEPONLJKONCELPQOLGEYKFNALLVRYTKVPQVST 420

DB 361 CATDPPACYAHVDFEPKPLVEEPNVLKTNCELPKELGEYGFQVALLVRYTKVPQVST 420

QY 421 PTLVEVSRLGKVGSKCKHPKAMPKCAEDYLSVVLNQLCVLHKEKTPVSDRVTKCCTES 480

DB 421 PTLVEVSRLGKVGSKCKTHPEARLSCAEDYLSVVLNRLCVLHKEKTPVSEKVTCKCTES 480

QY 481 LVNRRPCFSALVEVTYVPKFNARTFTFHADICTLSEKERQIKQKTALVELVGHKPKAT 540

DB 481 LVNRRPCFSALQVDVTYVPKFSFAETFTFHADLCTLPKAEQIKQKQSAVALLKHKPKAT 540

QY 541 KEQLKAVNMDFAAFVEKCKCKADDEKTCFAEKGKLVAAASQAAL 583

DB 541 EQLKTVNMGDFGFSFVDCAAEDKCAEFGPKLVAAQAAL 583

RESULT 6

ALBU CANFA

AC P49822; O7705; O9TSZ4; PRT; 608 AA.

DT 01-OCT-1996 (Rel. 34, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE Serum albumin precursor (Allergen Can f 3).

CN Name=ALB;

OS Canis familiaris (Dog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

OX NCBI\_TaxID=9615;

RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Beagle; TISSUE=Liver;  
RA Hilger C.;  
RL Submitted (MAR-1999) to the EMBL/GenBank/DBSJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=20148667; PubMed=10669848; DOI=10.1016/S0091-6749(00)90077-0;  
RA Pandjaitan B., Swoboda I., Brandejsky-Pichler F., Rumpold H.,  
VA Valenta R., Spitzauer S.;  
RT "Escherichia coli expression and purification of recombinant dog  
albumin, a cross-reactive animal allergen.";  
RL J. Allergy Clin. Immunol. 105:279-285(2000).  
RN [3]  
RP SEQUENCE OF 25-48.  
RX MEDLINE=75011422; PubMed=4414013;  
RA Dixon J.W., Sarkar B.;  
RT "Isolation, amino acid sequence and copper(II)-binding properties of  
peptide (1-24) of dog serum albumin.";  
RL J. Biol. Chem. 249:5872-5877(1974).  
RN [4]  
RP SEQUENCE OF 25-38.  
RC TISSUE=Heart;  
RX MEDLINE=98163340; PubMed=9504812;  
RA Dunn M.J., Corbett J.M., Wheeler C.H.;  
RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of  
dog heart proteins.";  
RL Electrophoresis 18:2795-2802(1997).  
RN [5]  
RP SEQUENCE OF 215-478 FROM N.A.  
RC TISSUE=Salivary gland;  
RX MEDLINE=94201492; PubMed=7512102;  
RA Spitzauer S., Schweiger C., Sperr W.R., Pandjaitan B., Valent P.,  
Muehl S., Ebner C., Scheiner O., Kraft D., Rumpold H.;  
RT "Molecular characterization of dog albumin as a cross-reactive  
allergen.";  
RL J. Allergy Clin. Immunol. 93:614-627(1994).  
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good  
binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
hormones, bilirubin and drugs. Its main function is the regulation  
of the colloidal osmotic pressure of blood.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Plasma.  
CC -!- ALLERGEN: Causes an allergic reaction in human.  
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.  
CC -!- SIMILARITY: Contains 3 albumin domains.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC EMBL; AJ133489; CAB64867.1; -;  
DR EMBL; Y17737; CAB76841.1; -;  
DR EMBL; S72946; AAB30434.1; -;  
DR HSSP; P02768; 1E7B.  
DR HSC-2DPAGE; P49822; DOG.  
DR InterPro; IPR000264; Serum albumin.  
DR Pfam; PF00273; Serum\_albumin; 3.  
DR PRINTS; PR00802; SERUMALBUMIN.  
DR ProDom; PD002486; Serum\_albumin; 1.  
DR SMART; SM00103; ALBUMIN; 3.  
DR PROSITE; PS00212; ALBUMIN; 3.  
KW Allergen; Copper; Direct protein sequencing; Lipid-binding;  
KW Metal-binding; Repeat; Signal  
FT SIGNAL 1 18 Potential.  
FT PROPEP 19 24  
FT PROPEP 25 608 Serum albumin.  
FT DOMAIN 25 205 Albumin 1.

```
FT DOMAIN 212 397 Alumin 2.
FT METAL 404 595 Copper (By similarity).
FT DISULFID 77 86 By similarity.
FT DISULFID 99 115 By similarity.
FT DISULFID 114 125 By similarity.
FT DISULFID 148 193 By similarity.
FT DISULFID 192 201 By similarity.
FT DISULFID 224 270 By similarity.
FT DISULFID 269 277 By similarity.
FT DISULFID 289 303 By similarity.
FT DISULFID 302 313 By similarity.
FT DISULFID 340 385 By similarity.
FT DISULFID 384 393 By similarity.
FT DISULFID 416 462 By similarity.
FT DISULFID 461 472 By similarity.
FT DISULFID 485 501 By similarity.
FT DISULFID 500 511 By similarity.
FT DISULFID 538 583 By similarity.
FT DISULFID 582 591 By similarity.
FT CONFLICT 1 26 MKWTFISLFFLFSSAYSRLVRREA -> MDT (in Ref. 2).
FT CONFLICT 146 146 A -> R (in Ref. 2).
FT CONFLICT 206 206 I -> T (in Ref. 2).
FT CONFLICT 349 349 V -> A (in Ref. 2).
FT CONFLICT 359 359 S -> A (in Ref. 1).
FT CONFLICT 448 448 V -> VV (in Ref. 1).
FT CONFLICT 474 474 D -> E (in Ref. 1).
SQ SEQUENCE 608 AA; 68606 MW; 3CFIC8FF7DD8FC06 CRC64;

Query Match 82.6%; Score 2562; DB 1; Length 608;
Best Local Similarity 79.8%; Pred. No. 8e-154;
Matches 465; Conservative 57; Mismatches 61; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFKDLGEBNFKALVLIAPQYLOQCPEFHVHKLNVETEFKTCVADESA 60
DB 25 EAYKEIAHRYNDLGEHFRGLVAFSOLQOCPEFHVHKLAKVETEFKAKAEESGA 84

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQPERNECFLOHKDNDPNLRLVRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQPERNECFLOHKDNDPNLRLVRPEV 144

QY 121 DVMCTAFHNEETFLKKYLYEIAARRHPYFYAPPELLFPAKRYKAAFTCECCQAADKAACLLP 180
DB 145 DALCAAFQDNEQLFGKLYEIAARRHPYFYAPPELLYAAQYKGVFAECCQAADKAACLGP 204

QY 181 KLDELDECKASSAKORLKCASLQKFGGERAFKAWAVARLSQRPKAEPAEVSCLVTDLT 240
DB 205 KIEALREKVLLSAKERFKCASLQKFGDRFAKAWAVARLSQRPKADFAEISKVTDLT 264

QY 241 VHTECHGDLLECADRADLAKYICENQDSISSKLKECCPELLEKSHCIAEVENDEMPA 300
DB 265 VHKECCGDLLECADRADLAKYICENQDSISSKLKECCDKPVLEKSCQLAEVERDELPG 324

QY 301 DLPSLAADPVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVVLLRLAKTYETTLK 360
DB 325 DLPSLAADPVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVVLLRLAKTYETTLK 384

QY 361 CAAADPHECYAKVDFEKLVEEPQNLKONCELPOLGEYKFONALLVRYTKVPQVST 420
DB 385 CATDDPPTCYAKVDFEKLVEEPQNLKONCELPOLGEYKFONALLVRYTKVPQVST 444

QY 421 PTLVEVSRNLGVSKCKKHPEAKRMPCAEDYLSVVNLQCVLHEKTPVSDRVTKCTSES 480
DB 445 PTLVEVSRNLGVSKCKKHPEAKRMPCAEDYLSVVNLQCVLHEKTPVSDRVTKCTSES 504

QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKQTALVELVKKHKPKAT 540
DB 505 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKQTALVELVKKHKPKAT 564

QY 541 KEQLKAVMDDFAAFVEKCKCKADDKTCFAEEGKLVAAASQAAL 583
DB 565 DEQLTKVMGDFGAFVEKCKCAENKEGCFSEEGKLVAAASQAAL 607
```

## RESULT 7

Q95VB7

```
ID Q95VB7 PRELIMINARY; PRT; 608 AA.
AC Q95VB7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Alumin.
OS Schistosoma mansoni (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC Schistosomatidae; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6183;
RN [1]
RP SEQUENCE FROM N.A.
RA Osman A., Asahi H., Staecker M.J., LoVerde P.T.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DDJ databases.
DR EMBL; AF418550; AAL08579.1; -.
DR HSSP; P02768; IHKL.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR001703; AlphaFetoprot.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; Serum_albumin; 3.
DR PRINTS; PRO0803; AFETOPROTEIN.
DR PRINTS; PRO0802; SERUMALBUMIN.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 608 AA; 68225 MW; E5EAB828E1C6E54 CRC64;
```

```
Query Match 80.7%; Score 2504; DB 2; Length 608;
Best Local Similarity 76.3%; Pred. No. 3.7e-150;
Matches 445; Conservative 79; Mismatches 59; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFKDLGEBNFKALVLIAPQYLOQCPEFHVHKLNVETEFKTCVADESA 60
DB 25 DAHKEIAHRYNDLGEHFRGLVAFSOLQOCPEFHVHKLNVETEFKTCVADESA 84

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQPERNECFLOHKDNDPNLRLVRPEV 120
DB 85 NCDKSLHTLFGDKLCAIPTLSDSYGELADCCAKQPERNECFLOHKDNDPNLPPFVRPDA 144

QY 121 DVMCTAFHNEETFLKKYLYEIAARRHPYFYAPPELLFPAKRYKAAFTCECCQAADKAACLLP 180
DB 145 EAMCTSFQENAVTTFMGHVLHEVARHPYFYAPPELLYAAEYSAIMTECCGEADKAACTP 204

QY 181 KLDELDECKASSAKORLKCASLQKFGGERAFKAWAVARLSQRPKAEPAEVSCLVTDLT 240
DB 205 KLDALKEKALASSVNRKLCSSLQRFQGRFAKAWAVARMSQRPKADFAEITKLATDLTK 264

QY 241 VHTECHGDLLECADRADLAKYICENQDSISSKLKECCPELLEKSHCIAEVENDEMPA 300
DB 265 LTBECCHGDLLECADRADLAKYICENQASISSKLQACCDKPVLSKSHCLSEVENDDLPA 324

QY 301 DLPSLAADPVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVVLLRLAKTYETTLK 360
DB 325 DLPSLAADPVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVVLLRLAKTYETTLK 384

QY 361 CAAADPHECYAKVDFEKLVEEPQNLKONCELPOLGEYKFONALLVRYTKVPQVST 420
DB 385 CAEADPSACYGVLFDEFQPLVEEPKLVKANCELFKLGEGYFQNALIVRYTKAPQVST 444

QY 421 PTLVEVSRNLGVSKCKKHPEAKRMPCAEDYLSVVNLQCVLHEKTPVSDRVTKCTSES 480
DB 445 PTLVEVSRNLGVSKCKKHPEAKRMPCAEDYLSVVNLQCVLHEKTPVSDRVTKCTSES 504

QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKQTALVELVKKHKPKAT 540
DB 505 VVERPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKQTALVELVKKHKPKAT 564

QY 541 KEQLKAVMDDFAAFVEKCKCKADDKTCFAEEGKLVAAASQAAL 583
```







Db 204 KIETREKVLASSARQRLRCASIQKFGERALKAMSVARLSQKPPKAEFVEVTKLVTDLT 263

Qy 241 VHTCECHGDLLECADDRLADLAKYICENQDSISSKKECCCKPLLEKSHCIAEVENDEMPA 300

Db 264 VHKCECHGDLLECADDRLADLAKYICDNQDTISSKKECCCKPLLEKSHCIAEVENDAIPE 323

Qy 301 DLPSLAADFEVSKDYCKNYAEAKDVLGMLFLEYARRHPDYSVLLLRLLAKTYETLEKC 360

Db 324 NLPLTADFAEDKDYCKNYQEAADFLGSLFLEYARRHPDYAVSVLLRLAKEYEATLEBC 383

Qy 361 CAADAPHECYAKVDFEFLVPEPQNLIKONCELFEOQGEYKFNALLVRYTKVPOVST 420

Db 384 CAKDDPHACYATVDFKLKHLVDEPQNLIKONCELFEOQGEYKFNALLVRYTKVPOVST 443

Qy 421 PTLVEVSRNLGVKSGCKKHPEAKRMPCAEDYLSVNLQCVLHEKTPVSDRVTCKCTES 480

Db 444 PTLVEVSRNLGVKSGCKKHPEAKRMPCAEDYLSVNLQCVLHEKTPVSEKVTCKCTES 503

Qy 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKKHPKAT 540

Db 504 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKKHPKAT 563

Qy 541 KEOLKAVMDDFAAFEVKECKCADDKCTCFAEEGKLVAAASQAAL 583

Db 564 DEQLKTVMENFVAFVDDKCAADDDKEACFAVEGPKLVVSTQTAL 606

RESULT 11

ALBU SHEEP

ID ALBU SHEEP STANDARD; PRT; 607 AA.

AC P14639;

DT 01-APR-1990 (Rel. 14, Created)

DT 01-APR-1990 (Rel. 14, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE Serum albumin precursor.

GN Name=ALB;

OS Ovis aries (Sheep).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Caprinae; Ovis.

OX NCBI\_TaxID=9940;

RN [1]\_TaxID=9940;

RP SEQUENCE FROM N.A.

RC TISSUE=Liver;

RX MEDLINE=90098888; PubMed=2602160;

RA Brown W.M., Dziegielewska K.M., Foreman R.C., Saunders N.R.;

RT "Nucleotide and deduced amino acid sequence of sheep serum albumin.";

RL Nucleic Acids Res. 17:10495-10495(1989).

CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good binding capacity for water, Ca(2+), Na(+), K(+), fatty acids, hormones, bilirubin and drugs. Its main function is the regulation of the colloidal osmotic pressure of blood.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- TISSUE SPECIFICITY: Plasma.

CC -!- SIMILARITY: Belongs to the ALB/AFB/VDB family.

CC -!- SIMILARITY: Contains 3 albumin domains.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; X17055; CAA34903.1; -.

DR PIR; S06936; A8SHS.

DR HSSP; P02768; 1HK1.

DR InterPro; IPR001703; Alphafetoprot.

DR InterPro; IPR000264; Serum albumin.

DR Pfam; PF00273; Serum albumin; 3.

DR PRINTS; PR00802; SERUMALBUMIN.

DR ProDom; PD002486; Serum\_albumin; 1.

DR SMART; SM00103; ALBUMIN; 3.

DR PROSITE; PS00212; ALBUMIN; 3.

KW Copper; Lipid-binding; Metal-binding; Repeat; Signal.

FT SIGNAL 1 18

FT PROPEP 19 24

FT CHAIN 25 607

FT DOMAIN 25 204

FT DOMAIN 211 396

FT DOMAIN 403 594

FT METAL 27 27

FT DISULFID 77 86

FT DISULFID 99 115

FT DISULFID 114 125

FT DISULFID 147 192

FT DISULFID 191 200

FT DISULFID 223 269

FT DISULFID 268 276

FT DISULFID 288 302

FT DISULFID 301 312

FT DISULFID 339 384

FT DISULFID 383 392

FT DISULFID 415 461

FT DISULFID 460 471

FT DISULFID 484 500

FT DISULFID 499 510

FT DISULFID 537 582

FT DISULFID 581 590

SQ SEQUENCE 607 AA; 69188 MW; 84979A87F8B86596 CRC64;

Query Match 78.4%; Score 2432.5; DB 1; Length 607;

Best Local Similarity 75.0%; Pred. No. 1.2e-145;

Matches 437; Conservative 73; Mismatches 72; Indels 1; Gaps 1;

Qy 1 DAHSEVAHRKFDLGEENFKALVLIATAQYLQCCPFDEHVKLVNEVTEFAKTCVADESAB 60

Db 25 DTHKSEIAHRFNDLGEENFQGLVLIATFQYLQCCPFDEHVKLVKELTEFAKTCVADESAB 84

Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 120

Db 85 GCDKSLHTLFGDELCKVATLRETYGDMADCCQEPERNECFLNHKDDSPDLKPKL-KPEP 143

Qy 121 DVMCTAFHDNEETFLKLYEIAIRHPYFABELLFFAKRYKAAFTCECCOAAADKACLLP 180

Db 144 DTLCAEFKADKKFWGKYLYEVARHPYFABELLYANKYNGVFQEQCAEDKAGACLLP 203

Qy 181 KLDELDEGKASSAKORLKCASLOKFGGERAFKAWAVARLSQRFKAEFAEVSGLVTDLT 240

Db 204 KIDANREKVLASSARQRLRCASIQKFGERALKAMSVARLSQKPPKADFTDVTKIYTDLT 263

Qy 241 VHTCECHGDLLECADDRLADLAKYICENQDSISSKKECCCKPLLEKSHCIAEVENDEMPA 300

Db 264 VHKCECHGDLLECADDRLADLAKYICDNQDTISSKKECCCKPLLEKSHCIAEVENDAIPE 323

Qy 301 DLPSLAADFEVSKDYCKNYAEAKDVLGMLFLEYARRHPDYSVLLLRLLAKTYETLEKC 360

Db 324 NLPLTADFAEDKDYCKNYQEAADFLGSLFLEYARRHPDYAVSVLLRLAKEYEATLEBC 383

Qy 361 CAADAPHECYAKVDFEFLVPEPQNLIKONCELFEOQGEYKFNALLVRYTKVPOVST 420

Db 384 CAKDDPHACYATVDFKLKHLVDEPQNLIKONCELFEOQGEYKFNALLVRYTKVPOVST 443

Qy 421 PTLVEVSRNLGVKSGCKKHPEAKRMPCAEDYLSVNLQCVLHEKTPVSDRVTCKCTES 480

Db 444 PTLVEVSRNLGVKSGCKKHPEAKRMPCAEDYLSVNLQCVLHEKTPVSEKVTCKCTES 503

Qy 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKKHPKAT 540

Db 504 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKKHPKAT 563

Qy 541 KEOLKAVMDDFAAFEVKECKCADDKCTCFAEEGKLVAAASQAAL 583

Db 564 DEQLKTVMENFVAFVDDKCAADDDKEACFAVEGPKLVVSTQTAL 606

```

RESULT 12
ALBU RAT
ID ALBU RAT STANDARD; PRT; 608 AA.
AC P02770; P11382;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Serum albumin precursor [Contains: Neurotensin-related peptide (NRP)].
GN Name=Alb;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81223722; PubMed=7017712;
RA Sargent T.D., Yang M., Bonner J.;
RT "Nucleotide sequence of cloned rat serum albumin messenger RNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 78:243-246(1981).
RN [2]
RP SEQUENCE OF 1-38, AND PROCESSING.
RX MEDLINE=77249657; PubMed=893447;
RA Strauss A.W., Bennett C.D., Donohue A.M., Rodkey J.A., Alberts A.W.;
RT "Rat liver pre-proalbumin: complete amino acid sequence of the pre-
RT piece. Analysis of the direct translation product of albumin messenger
RT RNA.";
RL J. Biol. Chem. 252:6846-6855(1977).
RN [3]
RP SEQUENCE OF 25-222.
RX MEDLINE=78109429; PubMed=564345;
RA Isemura S., Ikenaka T.;
RT "Amino acid sequences of fragments I and II obtained by cyanogen
RT bromide cleavage of rat serum albumin.";
RL J. Biochem. 83:35-48(1978).
RN [4]
RP SEQUENCE OF 223-288, AND 572-608.
RX MEDLINE=76260153; PubMed=956149;
RA Isemura S., Ikenaka T.;
RT "Fragmentation of rat serum albumin by cyanogen bromide cleavage and
RT the amino acid sequences of four fragments.";
RL J. Biochem. 79:1183-1196(1976).
RN [5]
RP SEQUENCE OF 166-174.
RX TISSUE=Plasma;
RX MEDLINE=81194805; PubMed=2437111;
RA Caraway R.E., Mitra S.P., Cochran D.E.;
RT "Structure of a biologically active neurotensin-related peptide
RT obtained from pepsin-treated albumin(s).";
RL J. Biol. Chem. 262:5968-5973(1987).
RN [6]
RP COPPER-BINDING.
RX MEDLINE=79001617; PubMed=80265;
RA Aoyagi Y., Ikenaka T., Ichida F.;
RT "Copper(II)-binding ability of human alpha-fetoprotein.";
RL Cancer Res. 38:3483-3486(1978).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- FUNCTION: NRP regulates fat digestion, lipid absorption, and blood
CC flow (potential).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/)
CC -----

```

```

or send an email to license@isb-sib.ch.
-----
CC EMBL; V01222; CAA24532.1; -.
CC PIR; A93872; ABRIS.
CC HSSP; P02768; 1HK1.
CC Rat-heart-2DPAGE; P02770; -.
CC RGD; 2085; Alb.
CC InterPro; IPR001703; Alphafetoprot.
CC InterPro; IPR000264; Serum albumin.
CC Pfam; PF00273; Serum_albumin; 3.
CC PRINTS; PR00802; SERUMALBUMIN.
CC ProDom; PD002486; Serum albumin; 1.
CC SMART; SM00103; ALBUMIN; 3.
CC PROSITE; PS00212; ALBUMIN; 3.
CC Copper; direct protein sequencing; Lipid-binding; Metal-binding;
CC Repeat; Signal.
CC SIGNAL 1 18
CC PROPEP 19 24
CC CHAIN 25 608
CC PEPTIDE 166 174
CC DOMAIN 25 205
CC DOMAIN 212 397
CC DOMAIN 404 595
CC METAL 27 27
CC DISULFID 77 86
CC DISULFID 99 115
CC DISULFID 114 125
CC DISULFID 148 193
CC DISULFID 192 201
CC DISULFID 224 270
CC DISULFID 269 277
CC DISULFID 289 303
CC DISULFID 302 313
CC DISULFID 340 385
CC DISULFID 384 393
CC DISULFID 416 462
CC DISULFID 461 472
CC DISULFID 485 501
CC DISULFID 500 511
CC DISULFID 538 583
CC DISULFID 582 591
CC VARIANT 262 262
CC CONFLICT 174 174
CC SEQUENCE 608 AA; 68718 MW; 58B497A282411AB7 CRC64;
Query Match 78.2%; Score 2426; DB 1; Length 608;
Best Local Similarity 73.4%; Pred. No. 3.2e-145;
Matches 428; Conservative 82; Mismatches 73; Indels 0; Gaps 0;
QY 1 DAKHSEVAHRPKDLGEENFKALVLIAPAOYLQOCPPFEDHVKLVNEVTEFAKTCVADESA 60
Db 25 EAHKSEIAHRPKDLGEGHFKGLVLIAPSOYLQKCPYBEHILKLVQEVTDFAKTCVADENAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCAKOEPRNECFLOHKDDNPMLPRLVREPV 120
Db 85 NCDKSIHTLFGDKLCAIPKLDNYGELADCAKOEPRNECFLOHKDDNPMLPQFQDEA 144
QY 121 DVMTCTAHDEETELKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQADKAACLTP 180
Db 145 EAMCTSFQENPTSLFGLHYLHEVARHPYFYAPPELLFYAEKYNEVLTCQCTESDKAACLTP 204
QY 181 KLDELREDEGKASSAKQRLKCAKSLQKFGERAFAKAWARLSQRFKAEPAEYSLVTLDTLK 240
Db 205 KLDVKEKALVAAVRQRMKSSMORFGERAFKANAVARMQSRFFNAEPAEITKLATDTVK 264
QY 241 VHTCCCHGDLLEACDADRDLAKYTCENQDSISSKLKCECKEPLLEKSHICIAEVENDENPA 300
Db 265 INKECCHGDLLEACDADRDLAKYTCENQATISSKLQACCDKPVQLQKSQLAEETHDNI 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLVEYARHPDYVSVLLLRLLAKTYETTLK 360
Db 325 DLPSIAADFVEDKEVCKNYAEAKDVLGTFLYEYRRHPDYVSVLLLRLLAKTYETTLK 384

```



QY 361 CAADPHCYAKVDFEKLPLVEBPONLIKONCELFQOLGEYKFNALLVRYTKVQVST 420  
Db 385 CABGDPACVGTVLAEPFLVEPKLVKTNCELYEKLGEYGFQNALVRYTKAPQVST 444  
QY 421 PTLVEVSRNLGKVGSKCCXHPKAKRMPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCCTES 480  
Db 445 PTLVEAARNLGRVGTCKCTLPEAQRLPCEVEDYLSALNRLCVLHEKTPVSEKVTCKCSGS 504  
QY 481 LVNRRPCFSALEVDETVYVPKEFNAETFTPHADICTLSEKERQIKKOTALVELVKKHKPKAT 540  
Db 505 LVRRPCFSALTVDETVYVPKEFAETFTPHSDICTLPLDKEKQIKKOTALAELVKKHKPKAT 564  
QY 541 KEOLKAVMDPFAAFVEKCKCCKADDKETCFAEEGKLVAAQOAL 583  
Db 565 EDQLKTVMGDFAQFVDKCKCAADKDNCFATEGPNLVARSKEAL 607

RESULT 13  
ALBU\_PIG STANDARD; PRT; 605 AA.

AC P08835; Q29018;  
DT 01-NOV-1988 (Rel. 09, Created)  
DT 01-NOV-1988 (Rel. 09, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Serum albumin precursor (Fragment).  
GN Name=ALB;  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=89016582; PubMed=3174440;  
RA Baldwin G.S.; Weinstock J.;  
RT "Nucleotide sequence of porcine liver albumin.";  
RL Nucleic Acids Res. 16:9045-9045(1988).  
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good  
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
CC hormones, bilirubin and drugs. Its main function is the regulation  
CC of the colloidal osmotic pressure of blood.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Plasma.  
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.  
CC -!- SIMILARITY: Contains 3 albumin domains.

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X12422; CAA30970.1; -;  
DR EMBL; M36787; AAA30988.1; -;  
DR PIR; S01382; ABPGS.  
DR HSSP; P02768; 1E7H.  
DR InterPro; IPR000264; Serum albumin.  
DR Pfam; PF00273; Serum albumin.3.  
DR PRINTS; PR00802; SERUMALBUMIN.  
DR ProDom; PD002486; Serum\_albumin; 1.  
DR SMART; SM00103; ALBUMIN; 3.  
DR PROSITE; PS00212; ALBUMIN; 3.  
KW Copper; Lipid-binding; Metal-binding; Repeat; Signal.  
FT NON\_TER 1 1  
FT SIGNAL <1 16 By similarity.  
FT PROPEP 17 22 By similarity.  
FT CHAIN 23 605 Serum albumin.  
FT DOMAIN 23 202 Albumin 1.  
FT DOMAIN 209 394 Albumin 2.  
FT DOMAIN 401 592 Albumin 3.  
FT METAL 31 31 Copper (By similarity).

FT DISULFID 75 84 By similarity.  
FT DISULFID 97 113 By similarity.  
FT DISULFID 112 123 By similarity.  
FT DISULFID 145 190 By similarity.  
FT DISULFID 189 198 By similarity.  
FT DISULFID 221 267 By similarity.  
FT DISULFID 266 274 By similarity.  
FT DISULFID 286 300 By similarity.  
FT DISULFID 299 310 By similarity.  
FT DISULFID 337 382 By similarity.  
FT DISULFID 381 390 By similarity.  
FT DISULFID 413 459 By similarity.  
FT DISULFID 458 469 By similarity.  
FT DISULFID 482 498 By similarity.  
FT DISULFID 497 508 By similarity.  
FT DISULFID 535 580 By similarity.  
FT DISULFID 579 588 By similarity.  
FT CONFLICT 562 562 E -> D (in Ref. 1; AAA30988).  
SQ SEQUENCE 605 AA; 69410 MW; 3E55680DDIA1FAFF CRC64;

Query Match 77.7%; Score 2411.5; DB 1; Length 605;  
Best Local Similarity 76.0%; Pred. No. 2.6e-144;  
Matches 438; Conservative 67; Mismatches 70; Indels 1; Gaps 1;

QY 1 DAHKEVAHRPKDILGEENFKALVLIATAFYLOQCPEFDHVKLVNEVTEFAKTCVADESAE 60  
Db 23 DTYKSEIAHRPKDLGEQVFKGLVLIATAFSQHLQCCPYEEHVKLVREVTEFAKTCVADESAE 82  
QY 61 NCCKSLHTLFGDKLCTVATRLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVVRPEV 120  
Db 83 NCCKSIHTLFGDKLCAIPSLREHYGDLADCCKEPERNECFLOHKNDNPDPKLP-KPDP 141  
QY 121 DVMCTAFHDNEETFLKLYEIAIRHPYPYAPPELLFFAKRYKAAFTCCCOAADKAAACLLP 180  
Db 142 VALCADFQDEQKFWGKLYEIAIRHPYPYAPPELLVYAIYKDVFSQCCQAADKAAACLLP 201  
QY 181 KLDELDEGKASSAKORLKASLOKFGERAFKAWAVARLSQRPKAEFAEVSCLVTDLT 240  
Db 202 KIEHLREKVLTSAAKORLKASIQKFGERAFKAWLARLSQRPKADFTTEISKIVTDLAK 261  
QY 241 VHTCECHGDLPCADDRADLAKYICENODSISSKLKECCPKLLEKSHCIAEVNDEMPA 300  
Db 262 VHKECHGDLLECADDRADLAKYICENQDTISTKLECCDKPLLEKSHCIAAKRDELPA 321  
QY 301 DLPSLAADPVESKDYCKNYAEAKDVFLGMFLYEYARRHPDYSVLLLRLLAKTYETTTLEKC 360  
Db 322 DLNPLEHDFVEDKEVKYKAEKDVFLGTFLYEYSRRHPDYSVLLLRLLAKTYETTTLEDC 381  
QY 361 CAAADPHCYAKVDFEKLPLVEBPONLIKONCELFQOLGEYKFNALLVRYTKVQVST 420  
Db 382 CAKEDPPACYATVDFKQPLVDEPKNLIKQNCLEFKLGEYGFQNALIVRYTKVQVST 441  
QY 421 PTLVEVSRNLGKVGSKCCXHPKAKRMPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCCTES 480  
Db 442 PTLVEAARNLGRVGTCKCTLPEAQRLPCEVEDYLSALNRLCVLHEKTPVSEKVTCKCSGS 501  
QY 481 LVNRRPCFSALEVDETVYVPKEFNAETFTPHADICTLSEKERQIKKOTALVELVKKHKPKAT 540  
Db 502 LVNRRPCFSALTVDETVYVPKEFAETFTPHADICTLPEDEKQIKKOTALVELVKKHKPHAT 561  
QY 541 KEOLKAVMDPFAAFVEKCKCCKADDKETCFAEEGKLV 576  
Db 562 EEQLRTVLGNFAAFVQKCAAPDHEACFAVEGPKFV 597

RESULT 14  
Q68NH7 PRELIMINARY; PRT; 607 AA.  
AC Q68NH7;  
DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE Albumin.

OS	Sus scrofa (Pig).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
NCBI_TaxID=9823;	
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Sun S., Deng J., Zhou Y., Lu J., Wu X.;
RT	"Porcine serum albumin gene."
RL	Submitted (JUN-2004) to the EMBL/GenBank/DBSJ databases.
DR	EMBL; AY663543; AAT98610.1; -
DR	InterPro; IPR000264; Serum albumin.
DR	Pfam; PF00273; Serum albumin; 3.
DR	PRINTS; PR00802; SERUMALBUMIN.
DR	ProDom; PD002486; Serum albumin; 1.
DR	SMART; SM00103; ALBUMIN; 3.
DR	PROSITE; PS00212; ALBUMIN; 3.
SQ	SEQUENCE 607 AA; 69691 MW; 788DA13543CA99D8 CRC64;
Query Match	
Best Local Similarity 77.5%; Score 2404.5; DB 2; Length 607;	
Matches 437; Conservative 67; Mismatches 71; Indels 1; Gaps 1;	
QY	1 DAHKSEVAHRFKDLGEENFKALVLIAPAYQLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB	25 DTYKSEIAHRRFKDLGEQYFKGLVLIASFQHLQCCPYEHHVKLVREVTEFAKTCVADESAAE 84
QY	61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOPEPERNECFLOHKDNNPRLVREV 120
DB	85 NCDKSLHTLFGDKLCAISLREHYGDLADCCKEPEPERNECFLOHKDNNPRLVREV 143
QY	121 DVMTAFHDNBEETFLKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCOAAADKAACLLP 180
DB	144 VALCADFOEDQKFWGKYLIEIARRHPYFYAPPELLFYAIYKDVSECCQAADKAACLLP 203
QY	181 KLDELRDGKASSAKQRLKASLOKFGGERAFKAWAVARLSQRPKAEFAVSKLVTDLTK 240
DB	204 KIEHLREKVLTSAAQRLKASIQKFGGERAFKANSRLRSQRPKADFTESKIVTDLAK 263
QY	241 VHTCCCHGDLLECCADRADLAKYICENODTSISSKLKCECKEPLEKSHCIAEAVENDMPA 300
DB	264 VHTCCCHGDLLECCADRADLAKYICENQDTISTKLKCECKEPLEKSHCIAEAKRDELPA 323
QY	301 DLPSLAADPFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVVLLLRLAKEYTETLEKC 360
DB	324 DLNPLEHDFVEDKEVKCKNYKAKHVLGTFLEYYSRRHPDYSVSLLLRIAKIYEATLEDK 383
QY	361 CAADPHCEYAKVDFEPLVEEPONLIKQNCLEPEQLGEYKFNALLVRYTKKVPQVST 420
DB	384 CAKEDPPACATVDFKQPLVDEPKNLIKQNCLEPEKLGEGYGFQNALVRYTKKVPQVST 443
QY	421 PTLVEVSRLGKVGSKCKKHPKAEKMPCAEDYLSVNLNOLCVLHEKTPVSDRVTKCCTES 480
DB	444 PTLVEVARLGLVGRSCCRPEERLSCAEDYLSVNLNRLCVLHEKTPVSEKVTCKCCTES 503
QY	481 LVNRRPCFSALVEVDYTVPKFENAFETFTFHADICTLSEKERQIKKQATLALVELVKKHKKPAT 540
DB	504 LVNRRPCFSALTPTDITYPKFEVECTFTFHADLCTLPDEKQIKKQATLALVELLKKHKKPAT 563
QY	541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLV 576
DB	564 BEQLRTVLGNFAAFVQKCAAPDHEACFAVEGPKFV 599
RESULT 15	
Q6WDN9	PRELIMINARY; PRT; 608 AA.
AC	Q6WDN9;
DT	05-JUL-2004 (TrEMBLrel. 27, Created)
DT	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE	Prealbumin precursor.
OS	Cavia porcellus (Guinea pig).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC	Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
NCBI_TaxID=10141;	
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Liver;
RA	Mihaylova-Todorova S., Choe S.M., Westfall D.P.;
RL	Submitted (APR-2003) to the EMBL/GenBank/DBSJ databases.
DR	EMBL; AY294645; AAO20088.1; -
DR	HSSP; P02768; IAO6
DR	GO; GO:0005615; C:extracellular space; IEA.
DR	GO; GO:0005386; F:carrier activity; IEA.
DR	GO; GO:0006810; P:transport; IEA.
DR	InterPro; IPR001703; Alphafetoprot.
DR	InterPro; IPR000264; Serum albumin.
DR	Pfam; PF00273; Serum albumin; 3.
DR	PRINTS; PR00803; AFETOPROTEIN.
DR	PRODOM; PR00802; SERUMALBUMIN.
DR	ProDom; PD002486; Serum albumin; 1.
DR	SMART; SM00103; ALBUMIN; 3.
DR	PROSITE; PS00212; ALBUMIN; 3.
KW	Signal.
FT	SIGNAL
FT	CHAIN
SQ	SEQUENCE 608 AA; 68889 MW; BBD510A78D0261BA CRC64;
Query Match	
Best Local Similarity 76.9%; Score 2387; DB 2; Length 608;	
Matches 421; Conservative 87; Mismatches 75; Indels 0; Gaps 0;	
QY	1 DAHKSEVAHRFKDLGEENFKALVLIAPAYQLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB	25 EAHKSEIAHRRFKDLGEGHFKGLVLIATLSQHLQKSPFEEHVKLVNEVTEFAKTCVADESAAE 84
QY	61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOPEPERNECFLOHKDNNPRLVREV 120
DB	85 NCGKAIATLFGDKVCAIPSLRETYGELADCCAKEDDPDRVECFLOHKDNNPRLVREV 144
QY	121 DVMTAFHDNBEETFLKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCOAAADKAACLLP 180
DB	145 EALCTAFKENDRIFGHLYEVSRRHPYFYAPPELLFYAEKYKNALTECCCAADKAACLLP 204
QY	181 KLDELRDGKASSAKQRLKASLOKFGGERAFKAWAVARLSQRPKAEFAVSKLVTDLTK 240
DB	205 KLDAIKEKALVSSAAQRLKASLOKFGGERAFKANSVARLSQRPKAEFAEISTIVTSITK 264
QY	241 VHTCCCHGDLLECCADRADLAKYICENODTSISSKLKCECKEPLEKSHCIAEAVENDMPA 300
DB	265 VTKCCCHGDLLECCADRADLAKYICENQDTSISSKLKCECKVPKPTLQKACILEIORDLEPT 324
QY	301 DLPSLAADPFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVVLLLRLAKEYTETLEKC 360
DB	325 ELPLAVDFVEDKEVKCKNYKAKHVLGTFLEYYSRRHPDYSIGMLLRIAKGYEAKLEKC 384
QY	361 CAADPHCEYAKVDFEPLVEEPONLIKQNCLEPEQLGEYKFNALLVRYTKKVPQVST 420
DB	385 CAEADPHACYAKVDFEPLVEEPONLIKQNCLEPEQLGEYKFNALLVRYTKKVPQVST 444
QY	421 PTLVEVSRLGKVGSKCKKHPKAEKMPCAEDYLSVNLNOLCVLHEKTPVSDRVTKCCTES 480
DB	445 PTLVEYARKLGVSGLTKCCSIPETERLSCTENYLAALINRLCILHEKTPVSEKVTCKCCTES 504
QY	481 LVNRRPCFSALVEVDYTVPKFENAFETFTFHADICTLSEKERQIKKQATLALVELVKKHKKPAT 540
DB	505 LVNRRPCFSALHVDYTVPKFFHADSFTFHADICTLPEKEQVKQKQALVELVKKHKKPAT 564
QY	541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAAASQAL 583
DB	565 BEQMKVTMGDFAAFLKKCCADNKEACFTEDGPKLVAKCQATL 607
RESULT 16	
ALBU_MERUN	STANDARD; PRT; 609 AA.
ID	ALBU_MERUN

AC O35090; (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Serum albumin precursor.  
 GN Name=ALB;  
 OS Meriones unguiculatus (Mongolian jird) (Mongolian gerbil).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Gerbillinae;  
 OC Meriones.  
 OX NCBI\_TaxID=10047;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NGS IDR; TISSUE=Liver;  
 RX MEDLINE=98116663; PubMed=9455485;  
 RA Yoshida K., Seto-Oshima A., Sinozawa H.;  
 RT "Sequencing of cDNA encoding serum albumin and its extrahepatic  
 synthesis in the Mongolian gerbil, Meriones unguiculatus.";  
 RL DNA Res. 4:351-354(1997).  
 CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good  
 binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
 hormones, bilirubin and drugs. Its main function is the regulation  
 of the colloidal osmotic pressure of blood.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Plasma.  
 CC -!- SIMILARITY: Belongs to the ALB/AFB/VDB family.  
 CC -!- SIMILARITY: Contains 3 albumin domains.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 CC EMBL; AB006197; BAA21765.1; --  
 DR FIR; JCS838; JCS838.  
 DR HSP; P02768; 1E7B.  
 DR InterPro; IPR001703; Alphafetoprot.  
 DR InterPro; IPR000264; Serum albumin.  
 DR Pfam; PF00273; Serum albumin; 3.  
 DR PRINTS; PR00802; SERUMALBUMIN.  
 DR ProDom; PD002486; Serum albumin; 1.  
 DR SMART; SM00103; ALBUMIN; 3.  
 DR PROSITE; PS00212; ALBUMIN; 3.  
 DR Copper; Lipid-binding; Metal-binding; Repeat; Signal.  
 FT SIGNAL 1 18 By similarity.  
 FT PROPEP 19 24 By similarity.  
 FT CHAIN 25 609 Serum albumin.  
 FT DOMAIN 25 206 Albumin 1.  
 FT DOMAIN 213 398 Albumin 2.  
 FT DOMAIN 405 596 Albumin 3.  
 FT METAL 28 28 Copper.  
 FT DISULFID 78 87 By similarity.  
 FT DISULFID 100 116 By similarity.  
 FT DISULFID 115 126 By similarity.  
 FT DISULFID 149 194 By similarity.  
 FT DISULFID 193 202 By similarity.  
 FT DISULFID 225 271 By similarity.  
 FT DISULFID 270 278 By similarity.  
 FT DISULFID 290 304 By similarity.  
 FT DISULFID 303 314 By similarity.  
 FT DISULFID 341 386 By similarity.  
 FT DISULFID 385 394 By similarity.  
 FT DISULFID 417 463 By similarity.  
 FT DISULFID 462 473 By similarity.  
 FT DISULFID 486 502 By similarity.  
 FT DISULFID 501 512 By similarity.  
 FT DISULFID 539 584 By similarity.  
 FT DISULFID 583 592 By similarity.  
 CC SEQUENCE 609 AA; 68940 MW; 9CASF97F67EF1A48 CRC64;

Query Match 76.9%; Score 2387; DB 1; Length 609;  
 Best Local Similarity 73.9%; Pred. No. 9.3e-143;  
 Matches 430; Conservative 65; Mismatches 187; Indels 0; Gaps 0;  
 QY 2 AHKSEVAHFRKDLGSENFKALVLIFAQVLOQCPPEDEHVKLVNEVTEFAKTCVADESAN 61  
 DB 27 AHKSEIAHRYKDLGEKYEKGLVLYTFSQYLQKCSYEEHVHVKLVREVTDFASNCAKDESAEN 86  
 QY 62 CDKSLHTLFGDKLCTVATLRTYTGEMADCCAKQEPERNECFLOHKDDNNPRLVRPEVD 121  
 DB 87 CDKSLHTLFGDKLCSLPNFGKYEAMADCCAKQEPERNECFLOHKDDNPQLPFPKRAEPD 146  
 QY 122 VMCTAFHDNEETFLKKLYEYARRHPYFAPBELLFFAKRYKAAFPTECCOAAKACLLPK 181  
 DB 147 AMCTAFQENAEAFMGHYLHEVARRHPYFYGPPELLYLDKTYTAVLTTECCAADDKGACLTTPK 206  
 QY 182 LDLELDEGKASSAKORLKCASLOKQEGEFAKAWAVARLSQRPKPAEAEVSKLVTDLTQV 241  
 DB 207 LDALKEKALVSAVRQRLKCSSMKKEGFAKAWAVARMSQTFPNADFAEITKLTADLTQV 266  
 QY 242 HTECHGDLLECADRADLAKYICNQDSISGKLKCEKPELLEKSHCIAEVENDEMPAD 301  
 DB 267 TOECCHGDLLECADRAELAKYMCNQASISGKLQACCDKEMLKQSQCLAEVEHDDMPAD 326  
 QY 302 LPSLAADFVESKQCKNYAEAKQVPLGMFLYAYARRHPDYSVVLRLAKTYETTLKCC 361  
 DB 327 LPALTADFVEDKQVCKNYAEAKQVPLGMFLYAYARRHPDYSVVLRLAKTYETTLKCC 386  
 QY 362 AAADPHECVAKYVDFEFKPLVEEPQNLIKONCELPQLGEYKFNQALLVRYTKKVPQVSTP 421  
 DB 387 AEADPHACYGHVDFEFKPLVEEPQNLVKSNCLEYKLGEGYFQNAVLYRYTKKAPQVSTP 446  
 QY 422 TLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHETKPTVSDRVTKCTESL 481  
 DB 447 TLVEAARSILGRVGHCCALPEKKRLPCVEDYLSAILNRVCLLHETKPTVSEQVTKCCSGSL 506  
 QY 482 VNRRCPFSALEVDETYVPKEFNAETFTPHADICTLSEKEROIKOTALVELVKKHPKATK 541  
 DB 507 VERRECFPSALPVDETYVPKEFNAETFTPHANICTLPEKEQKQEKQKOTALVELVKKHPQATE 566  
 QY 542 EQLKAVMDDFAAVFEKCKCKADDKCTCFABEGKKLVAASQAL 583  
 DB 567 EQLKVMGDFAEFLKCKCKQEDKEACFTGEPKLVAESQKAL 608  
 RESULT 17  
 QGB3Z0  
 ID Q6B3Z0 PRELIMINARY; PRT; 583 AA.  
 AC Q6B3Z0;  
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE Serum albumin (Fragment).  
 GN Name=ESA;  
 OS Elephas maximus (Indian elephant).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Proboscidea; Elephantidae; Elephas.  
 OX NCBI\_TaxID=9783;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lazar J., Rasmussen B., Greenwood D.R., Bang I.-S., Prestwich G.D.;  
 RT "Elephant Albumin: A Multi-purpose Pheromone Shuttle.";  
 RL Chem. Biol. 0:0-0(2004).  
 DR EMBL; AY684122; AAT90502.1;  
 DR GO; GO:0005615; C:extracellular space; IEA.  
 DR GO; GO:0005386; F:carrier activity; IEA.  
 DR GO; GO:0006810; P:transport; IEA.  
 DR InterPro; IPR000264; Serum albumin.  
 DR Pfam; PF00273; Serum albumin; 3.  
 DR PRINTS; PR00802; SERUMALBUMIN.  
 DR ProDom; PD002486; Serum albumin; 1.  
 DR SMART; SM00103; ALBUMIN; 3.  
 DR PROSITE; PS00212; ALBUMIN; 3.

```
FT NON TER 1
SQ SEQUENCE 583 AA; 65767 MW; 4EC031C2EC8B9141 CRC64;

Query Match 76.7%; Score 2379.5; DB 2; Length 583;
Best Local Similarity 73.7%; Pred. No. 2.6e-142;
Matches 429; Conservative 67; Mismatches 85; Indels 1; Gaps 1;

QY 3 HKSEVAHFKDLGEENFKALVLIAPQYLQCCPFEDHVKLVNVEVTEPAKTCVADESANEC 62
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 YKSEIAHYKDLGELDFGLLISFAQLQKSPYDEHVQSVTAVYDLAKTCADESABHC 60
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 63 DKSLHTLFGDKLCT-VATLRETYGMAWCCAKQEPERNECFLOHKDDNPFLVLRVPEVD 121
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 GDSLHTIFGDKLCARVTAHQDTYGFECACGQEPERNECFLOHKDDNPFLVLRVPPAD 120
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 122 VMCATFHDNEETFLKKLYETARRHPYFAPPELLFPAKRYKAAFTTECCQADKKAACLLPK 181
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 ALCAFSFEENERKFFGVLYEYARRHPYFAPPELIYSEKYKDLTECCHADDKAACLPK 180
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 182 LDLEDECKGASAKORLAKCASLOKFGERAFAKAWARLSQRPFAEVSGLVTLTKV 241
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 IDALKDVTVLSAARQLKCANIHKEGERAFKAWATHLSQRPFPKADFAEVSKLATDLGKV 240
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 242 HTECHGDLLECADRADLAKYICENQDSISSKLKECCKEPLLEKSHCIAEVENDEMPAD 301
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 241 YQECCHGDLLECADRADLAKYICDNQETLSSKLKECCDKPVLAKSHCIAELDKDDPPAD 300
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 302 LPSLAADFVESKVCNKYAEAKDVPFLGMLFVEYARRHPDYSVLLLLRLAKYETILEKCC 361
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 301 LPSIVPDVDEKDKVCKNYQEAQIDFLGFLYESSRHPESYAFLLRLTAKEYEATLERCC 360
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 362 AAADPHECVAKVDFEFKPLVEEPQNLIKONCELPOLGEYKFNALLVRYTKVPQVSTP 421
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 361 AAADPHACYAKVLDAFTFLVEEPQGLVKQNEHFQGLGEYFTNELAVRYTKVPQVSTP 420
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 422 TLVEYSRNLGKVGSKCKHPKAKRMPCAEDYLSVLNLQCLVHLHKTVPDSRVTKCCTESL 481
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 421 TLVEYSRNLGKVGSKCKLAERKMPCAEDYLSVLNLQCLVHLHKTVPDSRVTKCCTESL 480
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 482 VNRRCPFALEVDYEVYKFNFAETFTFHADICTLSEKRIKQTALVELVKHKPKATK 541
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 481 VNRRCPFTALQDTYVYKFNFAETFTFHADICTLPEEQKQIKQSIVLVELVKHKPKATD 540
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 542 EQLKAVMDDFAAFEVKCKCADDKETCFABEGKLVAAASOAL 583
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 541 EQLKVTVTQFTGWKEKCAAEDEKEACFAEBEGKLVITASOAL 582
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 18
ALBU MOUSE
ID _ALBU MOUSE STANDARD; PRT; 608 AA.
AC P07724; Q61802;
DT 01-APR-1988 (Rel. 07, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Serum albumin precursor.
GN Name=Alb; Synonyms=Alb-1, Alb1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA van Reeth T., Gabant P., Dreze P., Szpirer J., Szpirer C.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=22354583; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaïdo I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
```

```
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrester A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmer S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrowsky N., Pullai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tonita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
[3]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Kidney, and Liver;
MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Fellings E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raba S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Richards S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Bosak S.A., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smal M.A.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[4]
RP SEQUENCE OF 99-516 FROM N.A.
RX MEDLINE=98216123; PubMed=2452956;
RA Minghetti P.P., Law S.W., Dugaiczky A.;
RT "The rate of molecular evolution of alpha-fetoprotein approaches that
RT of pseudogenes.";
RL Mol. Biol. Evol. 2:347-358(1985).
[5]
RP SEQUENCE OF 477-551 FROM N.A.
RX STRAIN=BALB/c;
MEDLINE=90269606; PubMed=1971802; DOI=10.1016/0378-1119(90)90030-U;
RA Boccaccio C., Deschatrette J., Meunier-Rotival M.;
RT "Empty and occupied insertion site of the truncated LINE-1 repeat
RT located in the mouse serum albumin-encoding gene.";
RL Gene 88:181-186(1990).
[6]
RP SEQUENCE OF 25-44.
RC TISSUE=Liver;
RX MEDLINE=93162044; PubMed=1286668;
RA Giometti C.S., Taylor J., Tollakson S.L.;
RT "Mouse liver protein database: a catalog of proteins detected by two-
RT dimensional gel electrophoresis.";
RL Electrophoresis 13:970-991(1992).
```

CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good  
 CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
 CC hormones, bilirubin and drugs. Its main function is the regulation  
 CC of the colloidal osmotic pressure of blood.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Plasma.  
 CC -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.  
 CC -!- SIMILARITY: Contains 3 albumin domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AJ011413; CAA09617.1; -;  
 CC EMBL: AK010025; BAB26650.1; -;  
 CC EMBL: BC024643; AAH24643.1; -;  
 CC EMBL: BC049971; AAH49971.1; -;  
 CC EMBL: M16111; AAA37190.1; -;  
 CC EMBL: X13060; CAA31458.1; -;  
 CC PIR: A05139; A05139.  
 CC HSSP: P02768; LHK1.  
 CC SWISS-2DPAGE: P07724; MOUSE.  
 CC MGD: MGI:87991; Alb1.  
 CC InterPro: IPR001703; Alphafetoprot.  
 CC Pfam: PF00273; Serum albumin; 3.  
 CC PRINTS: P00802; SERUMALBUMIN.  
 CC ProDom: PD002486; Serum albumin; 1.  
 CC PROSITE: PS00212; ALBUMIN; 3.  
 CC COPSITE: Direct protein sequencing; Lipid-binding; Metal-binding;  
 KW Repeat; Signal.  
 FT SIGNAL 1 18 By similarity.  
 FT PROPEP 19 24  
 FT CHAIN 25 608 Serum albumin.  
 FT DOMAIN 25 205 Albumin 1.  
 FT DOMAIN 212 397 Albumin 2.  
 FT DOMAIN 404 595 Albumin 3.  
 FT METAL 27 27 Copper.  
 FT DISULFID 77 86 By similarity.  
 FT DISULFID 99 115 By similarity.  
 FT DISULFID 114 125 By similarity.  
 FT DISULFID 148 193 By similarity.  
 FT DISULFID 192 201 By similarity.  
 FT DISULFID 224 270 By similarity.  
 FT DISULFID 269 277 By similarity.  
 FT DISULFID 289 303 By similarity.  
 FT DISULFID 302 313 By similarity.  
 FT DISULFID 340 385 By similarity.  
 FT DISULFID 384 393 By similarity.  
 FT DISULFID 416 462 By similarity.  
 FT DISULFID 461 472 By similarity.  
 FT DISULFID 485 501 By similarity.  
 FT DISULFID 500 511 By similarity.  
 FT DISULFID 538 583 By similarity.  
 FT DISULFID 582 591 By similarity.  
 FT CONFLICT 27 27 H -> D (in Ref. 6).  
 FT CONFLICT 33 33 H -> D (in Ref. 6).  
 FT CONFLICT 41 41 Q -> I (in Ref. 6).  
 SQ SEQUENCE 608 AA; 68692 MW; 292F7C7EED3A61B4 CRC64;  
 Query Match 76.6%; Score 2378; DB 1; Length 608;  
 Best Local Similarity 72.4%; Pred. No. 3.4e-142;  
 Matches 422; Conservative 80; Mismatches 81; Indels 0; Gaps 0;  
 QY 1 DAKHSEVAHFRKDLGLENFALVIAFAQYLOQCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
 Db 25 EAHKSEIAHYNDLGSQHFHGLVIAFSQYLQKCSYDEHAKLVQEVYDFAKTCVADESAA 84  
 QY 61 NCDKSLHTLFGDKLCAINPLNRENYGELADCCCTQKQPERNECFLOHKDDNPNLPRVREPV 120

Db 85 NCDKSLHTLFGDKLCAINPLNRENYGELADCCCTQKQPERNECFLOHKDDNPNLPRVREPPA 144  
 QY 121 DVMCTAFADHNEETFLKLYEIAARRHPYFYAPPELLFFAKRYKAAFTTECCOAAADKAAACLLP 180  
 Db 145 EAMCTSFKENPTTFMGHLYHEVARRHPYFYAPPELLFYAEQYNEILLTQCCAEADKESCLTP 204  
 QY 181 KLDELURDEGKASSAKORLKCAQLQKFGERAFKAMAVARLSQFPKAEFAEVSKLVTDLTK 240  
 Db 205 KLDGVKEKALVSSVRQMKCSMQKFGERAFKAMAVARLSQTFPNADFAEITKLAIDLTK 264  
 QY 241 VHTCECHGDLLECADRADLAKYICENODSISSKLKECCCKPILLESKSHCIAEVENDEMPA 300  
 Db 265 VNKECHGDLLECADRADLAKYICENQATISSKLQTCCKPILLESKSHCIAEVEHDTMPA 324  
 QY 301 DLPSLAADFVESKDYCKNYAEAKDFLGMFLYEAARRHPDYSVLLRLAKVTETLEKC 360  
 Db 325 DLPALAAADFVEDQEVCKNYAEAKDFLGTFLYEAARRHPDYSVLLRLAKVTETLEKC 384  
 QY 361 CAAADPHCYAKVDFEFKPLVEEPQNLKQNCLEFQELGEYKFNALLVRYTKKVPQVST 420  
 Db 385 CAEANPPACYGTVLAEFQPLVEEPKLVKTNCDLYEKLGEYGFQNALVRYTKQKAPQVST 444  
 QY 421 PTLVEVSRLGKVGSKCKKHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480  
 Db 445 PTLVEAARNLGRVGTCKCTLPEDQRLPCVEDYLSAILNRVCLJLHEKTPVSEHVTKCCSGS 504  
 QY 481 LVNRRPCFSALEVDETVYPKEFNAETFTPHADICTLSEKEROIKKQATLVELVKKHKKPKAT 540  
 Db 505 LVERRPCFSALTVDETYVPKEFKATFTFHSIDICTLPEKEKQIKKQATLVELVKKHKKPKAT 564  
 QY 541 KEQKAVMDDFAAFVEKCKCKADDDKTCFAEBGKKLVAAQAAL 583  
 Db 565 ABQLKTVMDDFQAFLDTCCKAADKDTCTFSTEGPNLVTRCKDAL 607  
 RESULT 19  
 Q8C7H3 PRELIMINARY; PRT; 608 AA.  
 AC Q8C7H3;  
 DT 01-MAR-2003 (Tremblrel. 23, Created)  
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)  
 DE Mus musculus adult male liver tumor cDNA, RIKEN full-length enriched  
 DE library, clone:C73030P03 product:albumin 1, full insert  
 DE sequence.  
 GN Name=Alb1;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Liver;  
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
 RA Carninci P., Hayashizaki Y.;  
 RT "High-efficiency full-length cDNA cloning."  
 RL Meth. Enzymol. 303:19-44(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Liver;  
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
 RA RIKEN FANTOM Consortium;  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Liver;  
 RA The FANTOM Consortium;  
 RA The RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs."  
 RL Nature 420:563-573(2002).

```

RN RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Liver;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RL prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
[5]
RN RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Liver;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
[6]
RN RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Liver;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Kato H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tanaka T.,
RA Tagawa A., Takahashi F., Takaku-Akashira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toyota T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AK050248; BAC34145.1; -.
DR HSSP; P02768; 1HK1.
DR MGD; MGI:87991; Alb1.
DR GO; GO:0005615; C:extracellular space; TAS.
DR InterPro; IPR001703; Alphafetoprot.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; Serum albumin; 3.
DR PRINTS; PR00803; AFETOPROTEIN.
DR PRODOM; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
SQ SEQUENCE 608 AA; 68722 MW; 292F600EED3A61B4 CRC64;
Query Match 76.5%; Score 2374; DB 2; Length 608;
Best Local Similarity 72.2%; Pred. No. 6.2e-142;
Matches 421; Conservative 80; Mismatches 82; Indels 0; Gaps 0;
QY 1 DAHKSEVAHRPKDLGEENFKALVILAPAQYLQCCFFEDHVKLVNVEYFAKTCVADESAE 60
DB 25 EAHKSEIAHRYNDGEQHFKGVLAFSQYLQKCSYDEHAKLVQEVDTFAKTCVADESA 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPNECFLOHKDDNPRLVLRPEV 120
DB 85 NCDKSLHTLFGDKLCAIPNLRENYGELADCTCKQEPNECFLOHKDDNPSPPPERPEA 144
QY 121 DVMTCTAFHNBETFLKKYLYEATRRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180
DB 145 EAMCTSFENPTTFMGVHLHEVARHPYFYAPPELLYAEQYNEILTQCCEADKESCLTP 204
QY 181 KLDELREDEGKASSAKORLKCASLOKFGERAFKAWAVARLSQFPFAEFAEVSCLVTLTK 240
DB 205 KLDGVKEKALVSSVRKQSMQKFGERAFKAWAVARLSQTFPNADFAEITKATDLTK 264
QY 241 VHTCECHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300

```

```

DB 265 VNKECCCHGDLLECADRAELAKYCNENQATISSKLOTCCDKPLLKXAHCLSEVHDTPA 324
QY 301 DLPSLAADFVESKDVCKNXYABAKDVFGLMFELYEARHPDYSVVLLRLAKTYTTLEKC 360
DB 325 DLPAIAADFVEDQBSCKNXYAEAKDVFGLTFLYYSRRHPDYSVSLLLRLAKKYEATLEKC 384
QY 361 CAAADPHECYAKVDFEFPKPLVEEPONLIKONCELFEOALGEYKFNALLVRYTKVPOVST 420
DB 385 CAEANPPACYGTGLAEFQPLVEEPNVLKNCDDYELKGEYGFQNALVRYTKAPQVST 444
QY 421 PTLVEVSRLNKGKVGCKCKHPEAKRMPCABEDYLSVNLQCLVLEHKEFPVSDRVTKCCTES 480
DB 445 PTLVEAARNLGRVGTCKCTLPEDQRLFCVEDYLSAILNRVCLLHEKTPVSEHVTKCCSGS 504
QY 481 LVNRRPFCFSALEVEDTVVPKFEFNATETTFHADICTLSEKEROIKKQATLVELVHKHPKAT 540
DB 505 LVERRPFCFSAITVDETVVPKFEKATETTFHSDICTLPEKEKQIKKQATLALVELVHKHPKAT 564
QY 541 KEQLKAVMDPFAAFVEKCKKADDEKTCFAEKGKLVAAASQAAL 583
DB 565 AEQLKTVMDPFAQFLDTCCRAADKDCSTEGPNLVTRCKDTL 607

RESULT 20
Q8C7C7 Q8C7C7 PRELIMINARY; PRT; 576 AA.
AC Q8C7C7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Mus musculus 2 days neonate thymus thymic cells cDNA, RIKEN full-
DE length enriched library, clone:C920028B14 product:albumin 1, full
DE insert sequence. (Fragment).
GN Name=Albi;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
[3]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RA The FANTOM Consortium;
RT the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
[4]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
[5]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;

```





**THIS PAGE BLANK (USPTO)**

Seq 18 Fused with Higes4

- Fused

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2005, 08:25:53 ; Search time 167 Seconds  
(without alignments)  
1662.838 Million cell updates/sec

Title: US-10-933-523-18

Perfect score: 3785

Sequence: 1 DAHKEVAHRFKDLGEENFK.....leglireisgsvqlcyvhs 718

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.1

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%

Listing first 100 summaries

Database : A Geneseq\_16Dec04.\*

- 1: geneseqp1980s.\*
- 2: geneseqp1990s.\*
- 3: geneseqp2000s.\*
- 4: geneseqp2001s.\*
- 5: geneseqp2002s.\*
- 6: geneseqp2003as.\*
- 7: geneseqp2003bs.\*
- 8: geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3156.5	83.4	763	8	ADL16712 Human stu
2	3153.4	83.3	838	7	ADL16494 Human alb
3	3153.4	83.3	838	7	ADH21273 Human alb
4	3152.8	83.3	876	7	ADH21285 Human alb
5	3152.8	83.3	876	7	ADH21285 Human alb
6	3152.3	83.3	819	7	ADH21285 Human alb
7	3152.3	83.3	819	7	ADH21285 Human alb
8	3152.3	83.3	819	7	ADH21285 Human alb
9	3152.3	83.3	819	7	ADH21285 Human alb
10	3152.2	83.3	794	7	ADH21285 Human alb
11	3152.2	83.3	809	7	ADH21285 Human alb
12	3152.2	83.3	774	7	ADH21285 Human alb
13	3152.2	83.3	774	7	ADH21285 Human alb
14	3151	83.2	774	7	ADH21285 Human alb
15	3151	83.2	774	7	ADH21285 Human alb
16	3151	83.2	774	7	ADH21285 Human alb
17	3151	83.2	774	7	ADH21285 Human alb
18	3151	83.2	774	7	ADH21285 Human alb
19	3151	83.2	868	7	ADH21275 Human alb
20	3151	83.2	874	7	ADH21275 Human alb
21	3151	83.2	874	7	ADH21274 Human alb
22	3151	83.2	880	7	ADH21274 Human alb
23	3151	83.2	880	7	ADH21274 Human alb
24	3151	83.2	880	7	ADH21274 Human alb
25	3151	83.2	880	7	ADH21274 Human alb

99 3133.2 82.8 1010 7 ADF16432 Adf16432 Human alb  
100 3133.2 82.8 1010 7 ADF16431 Adf16431 Human alb

ALIGNMENTS

RESULT 1  
ADL16712  
ID ADL16712 standard; protein; 763 AA.  
XX  
AC ADL16712;  
XX  
DT 17-JUN-2004 (first entry)  
XX  
DE Human stumulatory factor-interleukin 11 (IL-11) fusion protein.  
XX  
KW haemostatic; antianaemic; nephrotropic; cytostatic; anti-HIV;  
KW immunosuppressive; cell proliferation stimulatory factor;  
KW human serum albumin; HSA; CPSP; haematological disorder; hypochromia;  
KW hypochromic microcytic anaemia; anaemia; Platelet-less; HIV infection;  
KW cancer; renal failure; tissue transplant; organ transplant; human;  
KW interleukin 11; IL-11; fusion protein.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
PN US2004063635-A1.  
XX  
PD 01-APR-2004.  
XX  
XX 26-JUN-2003; 2003US-00609346.  
XX  
PR 01-JUL-2002; 2002US-0392948P.  
XX  
PA (YUZZ/) YU Z.  
PA (FUYU/) FU Y.  
XX  
PI Yu Z, Fu Y;  
XX  
XX WPI; 2004-282521/26.  
DR N-PSDB; ADL16711.  
XX  
XX New isolated polynucleotide encoding a fusion protein formed between a  
PT human serum albumin (HSA) and a cell proliferation stimulatory factor  
PT (CPSP), useful for treating hematological disorders.  
XX  
PS Claim 34; SEQ ID NO 2; 65pp; English.  
XX  
CC The invention describes an isolated polynucleotide encoding a fusion  
CC protein formed between a human serum albumin (HSA) and a cell  
CC proliferation stimulatory factor (CPSP). Also described are: a  
CC recombinant vector comprising the isolated polynucleotide; a recombinant  
CC protein encoded by the polynucleotide; a recombinant cell comprising the  
CC recombinant vector; a composition comprising a combination of at least  
CC two different HSA/CPSP fusion proteins; a method for treating a patient  
CC with a CPSP, or hematological disorder; and a kit comprising a first  
CC fusion protein of HSA and a first CPSP, and a second fusion protein of  
CC HSA and a second CPSP. The polynucleotides, methods and compositions are  
CC useful for treating a patient needing CPSP or having haematological  
CC disorder, e.g. hypochromia, hypochromic microcytic anaemia and anaemia,  
CC platelet-less, HIV infection, cancer, renal failure and tissue/organ  
CC transplantation. This is the amino acid sequence of a fusion protein  
CC comprising human serum albumin (HSA) and cell stimulatory factor human  
CC interleukin 11 (IL-11).  
XX  
SQ Sequence 763 AA;

Query Match 83.4%; Score 3156.5; DB 8; Length 763;  
Best Local Similarity 80.9%; Pred. No. 5e-73;  
Matches 619; Conservative 11; Mismatches 40; Indels 95; Gaps 9;  
1 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60

Db 1 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
Qy 61 NCDKSLHTLFGDKLCTVATILRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 120  
Db 61 NCDKSLHTLFGDKLCTVATILRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 120  
Qy 121 DVNCTAFHDNEETFLKKYLYEYIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLLP 180  
Db 121 DVNCTAFHDNEETFLKKYLYEYIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLLP 180  
Qy 181 KLDELDEGKASSAKQRLKASLOKFGERAFAKAWAVARLSQRFPKAFBFAEVSKLVDLT 240  
Db 181 KLDELDEGKASSAKQRLKASLOKFGERAFAKAWAVARLSQRFPKAFBFAEVSKLVDLT 240  
Qy 241 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300  
Db 241 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300  
Qy 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYEVARRHPDYSVVLLRLAKTVETLEK 360  
Db 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYEVARRHPDYSVVLLRLAKTVETLEK 360  
Qy 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCLEFQOLGEYKFNALLVRYTKKVPQVST 420  
Db 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCLEFQOLGEYKFNALLVRYTKKVPQVST 420  
Qy 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
Db 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
Qy 481 LVNRRFCFSALEVDYTVPKFNAETFTPHADICTLSEKEROIKKOTALVELVKHKPKAT 540  
Db 481 LVNRRFCFSALEVDYTVPKFNAETFTPHADICTLSEKEROIKKOTALVELVKHKPKAT 540  
Qy 541 KEOLKAVMDDFAFVEKCKCKADDKETCFABEGKKLVAASQAALGLMSPRLVEPCSHALPQ 600  
Db 541 KEOLKAVMDDFAFVEKCKCKADDKETCFABEGKKLVAASQAALGLMSPRLVEPCSHALPQ 600  
Qy 601 G---LSP-----GVIIVRGLVLOEPKHFTVSLRQ----- 628  
Db 591 GPPRVSPDRAELDSTVLLTRSL-LADTRQLAAQLRDKFPADGDHNLDSLPTLAMSAGAL 649  
Qy 629 -AAHAPVTLRASPAD-----RTLQWISRWGOKKLIS----- 658  
Db 650 GALQLPGVLTRELRADLLSVLRHVQMLRRAGGSSLSKLTLEPELGTLOARLDRLRLQLLMS 709  
Qy 659 -----APPLFYPO-----RFPEVLLLFQEGGLKALN 685  
Db 710 RLALPQPPPPAPPAPPAPPSSAWGGIRAAHAIL---GGLHLTLTD 750  
RESULT 2  
ADFL16494  
ID ADFL16494 standard; protein; 838 AA.  
XX  
AC ADFL16494;  
XX  
DT 12-FEB-2004 (first entry)  
XX  
DE Human albumin therapeutic fusion protein SeqID1591.  
XX  
KW albumin fusion protein; albumin activity; human serum albumin;  
KW serum osmotic pressure; shelf-life; stability; antidiabetic;  
KW gene therapy; diabetes mellitus; human.  
XX  
OS Chimeric.  
OS Homo sapiens.  
XX  
PN WO2003060071-A2.  
PD 24-JUL-2003.  
XX

PP 23-DEC-2002; 2002WO-US040891.  
 XX 21-DEC-2001; 2001US-0341811P.  
 PR 24-JAN-2002; 2002US-0350358P.  
 PR 28-JAN-2002; 2002US-0351360P.  
 PR 26-FEB-2002; 2002US-0359370P.  
 PR 28-FEB-2002; 2002US-0360000P.  
 PR 27-MAR-2002; 2002US-0367500P.  
 PR 08-APR-2002; 2002US-0370227P.  
 PR 10-MAY-2002; 2002US-0378950P.  
 PR 24-MAY-2002; 2002US-0382617P.  
 PR 28-MAY-2002; 2002US-0383123P.  
 PR 05-JUN-2002; 2002US-0385708P.  
 PR 10-JUL-2002; 2002US-0394625P.  
 PR 24-JUL-2002; 2002US-0398089P.  
 PR 09-AUG-2002; 2002US-0402131P.  
 PR 13-AUG-2002; 2002US-0402708P.  
 PR 18-SEP-2002; 2002US-0411355P.  
 PR 18-SEP-2002; 2002US-0411426P.  
 PR 02-OCT-2002; 2002US-0414984P.  
 PR 11-OCT-2002; 2002US-0417611P.  
 PR 23-OCT-2002; 2002US-0420246P.  
 PR 05-NOV-2002; 2002US-0423623P.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA (DELZ) DELTA BIOTECHNOLOGY LTD.  
 PA (PRIN) PRINCIPIA PHARM CORP.  
 XX Ballance DJ, Turner AJ, Rosen CA, Haseltine WA;  
 XX WPI; 2003-598517/56.  
 XX New albumin fusion protein, useful for preparing a composition for  
 PT treating diabetes mellitus.  
 XX Example 4; SEQ ID NO 1591; 24pp; English.  
 XX This invention relates to a novel albumin fusion protein having albumin  
 CC or biological activity. Human serum albumin is responsible for a  
 CC significant proportion of the osmotic pressure of serum and also  
 CC functions as a carrier of endogenous and exogenous ligands. The fusion of  
 CC albumin to a therapeutic protein may increase shelf-life and stability of  
 CC the therapeutic protein. The albumin fusion protein of the invention may  
 CC allow production of compositions with antidiabetic activity whilst the  
 CC nucleotide sequence which encodes it may be useful for gene therapy. The  
 CC albumin fusion protein is useful for preparing a composition for treating  
 CC diabetes mellitus. The present sequence is the amino acid sequence of a  
 CC novel full-length human albumin therapeutic fusion protein of the  
 CC invention. Note: The sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/publishedpct\_sequences  
 XX Sequence 838 AA;  
 SQ  
 Query Match 83.3%; Score 3153.4; DB 7; Length 838;  
 Best Local Similarity 75.0%; Pred. No. 7e-73;  
 Matches 624; Conservative 15; Mismatches 37; Indels 156; Gaps 12;  
 QY 1 DAHKEVAHFRFDLGEENFKALVLAFAQYLOQCFEDHVKLVNVEVFAKTCVADESAAE 60  
 Db 25 DAHKEVAHFRFDLGEENFKALVLAFAQYLOQCFEDHVKLVNVEVFAKTCVADESAAE 84  
 QY 61 NCDKSLHTLFGDKLCTVATRLTYGEMADCCAKQPERNECFLOKHQNDNPNLPLVRPEV 120  
 Db 85 NCDKSLHTLFGDKLCTVATRLTYGEMADCCAKQPERNECFLOKHQNDNPNLPLVRPEV 144  
 QY 121 DVNCTAFHNDNEETFLKKLYELIARRHPYFAPPELLFFAKRYKAAFTCCQAADKAACLIP 180  
 Db 145 DVNCTAFHNDNEETFLKKLYELIARRHPYFAPPELLFFAKRYKAAFTCCQAADKAACLIP 204  
 QY 181 KLDELREDEKASAKORLKCASLQKFGGERAFKAWAVARLSQRFPAEVAEVSCLVDTLTK 240  
 Db 205 KLDELREDEKASAKORLKCASLQKFGGERAFKAWAVARLSQRFPAEVAEVSCLVDTLTK 264

QY 241 VHTCCCHGDLLECCADDRADLAKYICENQDSSISSKLKCCCKEPLEKSHCIAEVENDEMPA 300  
 Db 265 VHTCCCHGDLLECCADDRADLAKYICENQDSSISSKLKCCCKEPLEKSHCIAEVENDEMPA 324  
 QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYEVARRHPDYSVLLLRALKATYETTLK 360  
 Db 325 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYEVARRHPDYSVLLLRALKATYETTLK 384  
 QY 361 CAAADPHCYAKVDFEFPKPLVEEPPQLIKONCELFEOQLGEYKFNQALLVRYTKVPQVST 420  
 Db 385 CAAADPHCYAKVDFEFPKPLVEEPPQLIKONCELFEOQLGEYKFNQALLVRYTKVPQVST 444  
 QY 421 PTLVEVSRNLGKVGSKCKCKHPEAKRMPCAEADYLSVNLQCLVLEHKTVPVSDRVTKCCTES 480  
 Db 445 PTLVEVSRNLGKVGSKCKCKHPEAKRMPCAEADYLSVNLQCLVLEHKTVPVSDRVTKCCTES 504  
 QY 481 LVNRPPCFSALEVDVETVVPKEFNAETFEHADICTLSEKEROIKKQATALVELVHKPKAT 540  
 Db 505 LVNRPPCFSALEVDVETVVPKEFNAETFEHADICTLSEKEROIKKQATALVELVHKPKAT 564  
 QY 541 KEQLKAVMDDFAAFEVEKCKCKADDKETCFABEGKKLVAASQAALGLM----- 586  
 Db 565 KEQLKAVMDDFAAFEVEKCKCKADDKETCFABEGKKLVAASQAALGLMDLLQLFLFVLL 624  
 QY 587 -----SPRLEV----- 592  
 Db 625 SGMGATGLRTSLDPSLEIYKMKPEVKRREQLLAKNLAQLNDIHQQYKILDVMLKGLFK 684  
 QY 593 -----PCSHALPOGLSPGOVLIIVRGLVLOPKHFTVSLRDOAAHAPVTLRASPADR 643  
 Db 685 VLEDSRTVLTAAADVLPDGPFP-----QDEK-----LKDQAFSHV-VENTAFPGDV 727  
 QY 644 TLQ-----W--ISRWGQKLLISAPFLFYPQRFVLLLFQEGGLK--LALN 685  
 Db 728 VLRFPRIVHYFYFHNHNNLLIRWG---IS-----FCNQTVGVNQPHSPHLSILM 774  
 QY 686 GQGJG-----ATSMNQAL-----EQLREUR 706  
 Db 775 AQELGISEKSDNFQNPFKIDRTFIPSTDPFKALREBEKRRKKEKEKEIR 836  
 RESULT 3  
 ADH21793  
 ID ADH21793 standard; protein; 838 AA.  
 AC ADH21793;  
 XX  
 DT 11-MAR-2004 (first entry)  
 XX  
 DE Human albumin/HCE1P80 fusion protein, SEQ ID NO:590.  
 XX  
 KW Fusion protein; human serum albumin; HSA; therapeutic protein;  
 KW shelf-life; in vitro biological activity; in vivo biological activity;  
 KW metabolic disorder; endocrine disorder; diabetes; type 1; type 2;  
 KW diabetes-related condition; hyperglycaemia; neural disorder; neuropathy;  
 KW retinopathy; cardiovascular disorder; heart disease; renal disorder;  
 KW obesity; glucose level maintenance; weight loss; antidiabetic; cardiant;  
 KW anorectic; ophthalmological; gene therapy.  
 XX  
 OS Chimeric.  
 OS Homo sapiens.  
 XX  
 PN WO2003059934-A2.  
 XX  
 PD 24-JUL-2003.  
 XX  
 PF 23-DEC-2002; 2002WO-US040892.  
 XX  
 XX 21-DEC-2001; 2001US-0341811P.  
 PR 24-JAN-2002; 2002US-0350358P.  
 PR 26-FEB-2002; 2002US-0359370P.  
 PR 28-FEB-2002; 2002US-0360000P.  
 PR 28-FEB-2002; 2002US-0367500P.

```
PR 27-MAR-2002; 2002US-0367500P.
PR 08-APR-2002; 2002US-0370227P.
PR 10-MAY-2002; 2002US-0378950P.
PR 24-JUL-2002; 2002US-0398008P.
PR 03-AUG-2002; 2002US-0402131P.
PR 13-AUG-2002; 2002US-0402708P.
PR 18-SEP-2002; 2002US-0411355P.
PR 02-OCT-2002; 2002US-0414984P.
PR 11-OCT-2002; 2002US-0417611P.
PR 23-OCT-2002; 2002US-0420246P.
PR 05-NOV-2002; 2002US-0423623P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX PI Rosen CA, Haseltine WA;
XX WPI; 2003-598501/56.
XX
XX PT New albumin fusion protein, useful for preparing a composition for
XX PT treating diabetes mellitus.
XX
XX PS Disclosure; SEQ ID NO 590; 1086pp; English.
XX
XX CC The invention relates to fusion proteins comprising human serum albumin
XX CC (ADH2I530) and a therapeutic polypeptide such as a therapeutic protein,
XX CC antibody or peptide or their variants or fragments. The therapeutic
XX CC protein may be fused to the N-terminus, the C-terminus or both termini of
XX CC albumin via a linker. The albumin component of the fusion proteins
XX CC prolongs the shelf-life and the in vitro and vivo biological activity of
XX CC the proteins compared with those of the corresponding therapeutic
XX CC proteins on their own. The invention also relates to nucleic acids
XX CC encoding albumin fusion proteins, vectors and host cells comprising an
XX CC albumin fusion protein nucleic acid, compositions and kits comprising an
XX CC albumin fusion protein, the method of extending the shelf-life of a
XX CC therapeutic protein by fusion with albumin, and the treatment of disease
XX CC using an albumin fusion protein. The albumin fusion proteins may be used
XX CC in the treatment of metabolic/endocrine disorders, diabetes and diabetes-
XX CC related conditions. Specifically the albumin fusion proteins may be used
XX CC to treat type 1 and type 2 diabetes, hyperglycaemia, neural disorders
XX CC (especially neuropathy), retinopathy, cardiovascular disorders
XX CC (especially heart disease, renal disorders and obesity). The proteins may
XX CC also be used in a method of maintaining a basal glucose level in a
XX CC patient and in a method for losing weight. The present sequence is
XX CC related to the invention.
XX
XX SQ Sequence 838 AA;
Query Match 83.3%; Score 3153.4; DB 7; Length 838;
Best Local Similarity 75.0%; Pred. No. 7e-73;
Matches 624; Conservative 15; Mismatches 37; Indels 156; Gaps 12;
Qy 1 DAHKSEVAHRFKOLGEENFALVLIAPQYLOQCFEDHVKLVNEVTEFAKTCVADESAAE 60
Db 25 DAHKSEVAHRFKOLGEENFALVLIAPQYLOQCFEDHVKLVNEVTEFAKTCVADESAAE 84
Qy 61 NCDKSLHTLFGDKLCTVATRETYGEMADCCAKOEPERNECFLOHKDDNPNLRLVRPEV 120
Db 85 NCDKSLHTLFGDKLCTVATRETYGEMADCCAKOEPERNECFLOHKDDNPNLRLVRPEV 144
Qy 121 DVMTAFHDNEETFLKKYLYEIRRHRYFYFAPELLFFAKRYKAAFTCCQAADKAAACLLP 180
Db 145 DVMTAFHDNEETFLKKYLYEIRRHRYFYFAPELLFFAKRYKAAFTCCQAADKAAACLLP 204
Qy 181 KLDELREGKASSAKQRLKASLOKQGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 240
Db 205 KLDELREGKASSAKQRLKASLOKQGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 264
Qy 241 VHTCCGGDLLECCADRADLAKYICENQDSTSSKLECCCKPLEKSHCIAEVENDEMPA 300
Db 265 VHTCCGGDLLECCADRADLAKYICENQDSTSSKLECCCKPLEKSHCIAEVENDEMPA 324
Qy 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYEYARRHPDYSVVLLRLAKTYETTTLEKC 360
Db 325 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYEYARRHPDYSVVLLRLAKTYETTTLEKC 384
Qy 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELPQOLGEYKPFQNALLVRYTKKVPQVST 420
Db 385 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELPQOLGEYKPFQNALLVRYTKKVPQVST 444
Qy 421 PTLVEYSRNLGKVGSKCKKHPKAEKMPKAEEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 445 PTLVEYSRNLGKVGSKCKKHPKAEKMPKAEEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504
Qy 481 LVNRRPCFSALEVDYVYKPFNAETFTTHADICTLSEKEROIKKQOTALVELVKKHPKAT 540
Db 505 LVNRRPCFSALEVDYVYKPFNAETFTTHADICTLSEKEROIKKQOTALVELVKKHPKAT 564
Qy 541 KEOLKAVMDDEFAFVEKCKCKADDKETCFABEGKKLVAAASQAALGLM----- 586
Db 565 KEOLKAVMDDEFAFVEKCKCKADDKETCFABEGKKLVAAASQAALGLMDLLQLFLAFLFVLLL 624
Qy 587 -----SPRLV----- 592
Db 625 SGMGATGTLRTSLDPSLEIYKKMFVKKRREQLLAKNLAKLNDIHOQYKILDVMLKGLFK 684
Qy 593 -----PCSHALPOGLSPQGVIIIRGLVLOEPKHFTVSLRDQAAHAPVTLRASADR 643
Db 685 VLEDSRTVLTAAADVLPDGPFP-----QDEK-----LKDAFSHV-VENTAFPGDV 727
Qy 644 TLQ-----W--ISRWGOKKLISAPFLFPQPFVFLLLFOEGGLK--LALN 685
Db 728 VLRFPRIYHYFDHNSNNLLIRWG-----IS-----FCNQTGVFNQGGHSPILSLM 774
Qy 686 GQGLG-----ATSMNQOAL-----BQLREL 706
Db 775 AQELGISEKSNFQNFPKIDRTFIPSTDPFOKALAEKREKREKKEIR 826
RESULT 4
ADFL14940
ID ADF14940 standard; protein; 876 AA.
XX ADF14940;
AC ADF14940;
XX
XX DT 12-FEB-2004 (first entry)
XX
XX DE Human albumin therapeutic fusion protein SeqID236.
XX
XX KW albumin fusion protein; albumin activity; human serum albumin;
XX KW serum osmotic pressure; shelf-life; stability; antidiabetic;
XX KW gene therapy; diabetes mellitus; human.
XX
XX OS Chimeric.
XX OS Homo sapiens.
XX
XX PN WO2003060071-A2.
XX
XX PD 24-JUL-2003.
XX
XX PF 23-DEC-2002; 2002WO-US040891.
XX
XX PR 21-DEC-2001; 2001US-0341811P.
XX PR 24-JAN-2002; 2002US-0350358P.
XX PR 28-JAN-2002; 2002US-0351360P.
XX PR 26-FEB-2002; 2002US-0359370P.
XX PR 28-FEB-2002; 2002US-0360000P.
XX PR 27-MAR-2002; 2002US-0367500P.
XX PR 08-APR-2002; 2002US-0370227P.
XX PR 10-MAY-2002; 2002US-0378950P.
XX PR 24-MAY-2002; 2002US-0382617P.
XX PR 28-MAY-2002; 2002US-0383123P.
XX PR 05-JUN-2002; 2002US-0385708P.
XX PR 10-JUL-2002; 2002US-0394625P.
XX PR 24-JUL-2002; 2002US-0398008P.
XX PR 09-AUG-2002; 2002US-0402131P.
XX PR 13-AUG-2002; 2002US-0402708P.
```

PR 18-SEP-2002; 2002US-0411355P.  
 PR 18-SEP-2002; 2002US-0411426P.  
 PR 02-OCT-2002; 2002US-0414984P.  
 PR 11-OCT-2002; 2002US-0417611P.  
 PR 23-OCT-2002; 2002US-0420246P.  
 PR 05-NOV-2002; 2002US-0423623P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PA (DELZ) DELTA BIOTECHNOLOGY LTD.  
 PA (PRIN-) PRINCIPIA PHARM CORP.  
 XX  
 PI Ballance DJ, Turner AJ, Rosen CA, Haseltine WA;  
 XX  
 XX WPI; 2003-598517/56.  
 XX  
 DR New albumin fusion protein, useful for preparing a composition for  
 XX treating diabetes mellitus.  
 PT  
 XX  
 XX Example 4; SEQ ID NO 236; 24pp; English.  
 PS  
 XX This invention relates to a novel albumin fusion protein having albumin  
 CC or biological activity. Human serum albumin is responsible for a  
 CC significant proportion of the osmotic pressure of serum and also  
 CC functions as a carrier of endogenous and exogenous ligands. The fusion of  
 CC albumin to a therapeutic protein may increase shelf-life and stability of  
 CC the therapeutic protein. The albumin fusion protein of the invention may  
 CC allow production of compositions with antidiabetic activity whilst the  
 CC nucleotide sequence which encodes it may be useful for gene therapy. The  
 CC albumin fusion protein is useful for preparing a composition for treating  
 CC diabetes mellitus. The present sequence is the amino acid sequence of a  
 CC novel full-length human albumin therapeutic fusion protein of the  
 CC invention. Note: The sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/publishedpct\_sequences  
 XX  
 XX Sequence 876 AA;  
 SQ  
 Query Match 83.3%; Score 3152.8; DB 7; Length 876;  
 Beat Local Similarity 70.4%; Pred. No. 7.9e-73;  
 Matches 622; Conservative 13; Mismatches 36; Indels 212; Gaps 10;  
 QY 1 DAHKEVAHRFDKLGEEFNKALVLIATAQYLOQCFEDHVKLVNVTFAKTCVADESA 60  
 DB 25 DAHKEVAHRFDKLGEEFNKALVLIATAQYLOQCFEDHVKLVNVTFAKTCVADESA 84  
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPRLVLRPEV 120  
 DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPRLVLRPEV 144  
 QY 121 DVMCTAFHDNEETFLKLYETARRHPYFYAPPELLFFAKRYKAAFTCCQADKAACLLP 180  
 DB 145 DVMCTAFHDNEETFLKLYETARRHPYFYAPPELLFFAKRYKAAFTCCQADKAACLLP 204  
 QY 181 KLDELDEGKASSAKQRLKASLQKGFGERAFKAWAVARLSQRPKAEFAVSKLVTDLTK 240  
 DB 205 KLDELDEGKASSAKQRLKASLQKGFGERAFKAWAVARLSQRPKAEFAVSKLVTDLTK 264  
 QY 241 VTECHGDLLECADDRADLAKYICENQDSISLKECECEKPLEKSHCIAEVENDEMPA 300  
 DB 265 VTECHGDLLECADDRADLAKYICENQDSISLKECECEKPLEKSHCIAEVENDEMPA 324  
 QY 301 DLPLSAAADPVESKDVCKNVAEAKDVLGMFLVEYARRHPDYSVLLLLRLAKTYETTLK 360  
 DB 325 DLPLSAAADPVESKDVCKNVAEAKDVLGMFLVEYARRHPDYSVLLLLRLAKTYETTLK 384  
 QY 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELPQELGEYKFNALLVRYTKVPQVST 420  
 DB 385 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELPQELGEYKFNALLVRYTKVPQVST 444  
 QY 421 PTLVEVSRNLGVSKGCKCHGPEAKRMPCAEDVLSVLNQLCVLHEKTPVSDRVTTCCTES 480  
 DB 445 PTLVEVSRNLGVSKGCKCHGPEAKRMPCAEDVLSVLNQLCVLHEKTPVSDRVTTCCTES 504

QY 481 LVNRRPCFSALEVDYETVPKEFNAETTFHADICTLSEKERQIKKOTALVELVKHKPKAT 540  
 DB 505 LVNRRPCFSALEVDYETVPKEFNAETTFHADICTLSEKERQIKKOTALVELVKHKPKAT 564  
 QY 541 KEQLKAVMDDFAAFEVKECKKADDKETCFABEKGKLVAAASQAALGLMSPLREV 592  
 DB 565 KEQLKAVMDDFAAFEVKECKKADDKETCFABEKGKLVAAASQAALGL 620  
 QY 593 -----PCSHALPOGLSPGQVIVRGLVLQPKHPT----- 622  
 DB 621 RDAETGERLVCAQCPPTGTFVQRPCTRDSPITTCGP-----CPRHYTQFWMYLER 669  
 QY 623 -----VSLRDQAAHA----- 632  
 DB 670 CRYCNVLGGEREEARACHATHNRACRGTGFFAHAGFLEHASCPPGAGVIAFGTFSQN 729  
 QY 633 -----PVTLRAS----- 639  
 DB 730 TQCPQCPGTFSSASSSSSEQCPHNCALTALNVPGSSSHDTLCTCTGFLPLSTRVPG 789  
 QY 640 -----FADRTLQWISR-----WG-----QKLLISAPFLFPQ 666  
 DB 790 ABECERAVIDFVAFQDISIKRLQLLQALEAPEGWGPTPRAGRAALQKL-----R 840  
 QY 667 RPFVILLFQEGGLKLAINGOGLGATSNQOALEQLRELISG 709  
 DB 841 RRLTELLGAODGALLVRL-----LQALRVARMPPG 869  
 RESULT 5  
 ADH21285  
 ID ADH21285 standard; protein; 876 AA.  
 XX AC ADH21285;  
 XX  
 DT 11-MAR-2004 (first entry)  
 XX  
 XX Human albumin/mutant TR6 fragment fusion protein, SEQ ID NO:82.  
 DE  
 XX Fusion protein; human serum albumin; HSA; therapeutic protein;  
 KW shelf-life; in vitro biological activity; in vivo biological activity;  
 KW metabolic disorder; endocrine disorder; diabetes; type 1; type 2;  
 KW diabetes-related condition; hyperglycaemia; neural disorder; neuropathy;  
 KW retinopathy; cardiovascular disorder; heart disease; renal disorder;  
 KW obesity; glucose level maintenance; weight loss; antidiabetic; cardiant;  
 KW anorectic; ophthalmological; gene therapy.  
 XX  
 OS Synthetic.  
 OS Chimeric.  
 OS Homo sapiens.  
 XX  
 PN WO2003059934-A2.  
 XX  
 PD 24-JUL-2003.  
 XX  
 PF 23-DEC-2002; 2002WO-US040892.  
 XX  
 XX 21-DEC-2001; 2001US-0341811P.  
 PR 24-JAN-2002; 2002US-0350358P.  
 PR 26-FEB-2002; 2002US-0359370P.  
 PR 28-FEB-2002; 2002US-0360000P.  
 PR 27-MAR-2002; 2002US-0367500P.  
 PR 08-APR-2002; 2002US-0370227P.  
 PR 10-MAY-2002; 2002US-0378950P.  
 PR 24-JUL-2002; 2002US-0398008P.  
 PR 09-AUG-2002; 2002US-0402131P.  
 PR 13-AUG-2002; 2002US-0402708P.  
 PR 18-SEP-2002; 2002US-0411355P.  
 PR 02-OCT-2002; 2002US-0414984P.  
 PR 11-OCT-2002; 2002US-0417611P.  
 PR 23-OCT-2002; 2002US-0420246P.  
 PR 05-NOV-2002; 2002US-0423623P.  
 XX

PA	(HUMA-) HUMAN GENOME SCI INC.	
XX	Rosen CA, Haseltine WA;	
XX	WPI; 2003-598501/56.	
XX	New albumin fusion protein, useful for preparing a composition for	
PT	treating diabetes mellitus.	
XX	Disclosure; SEQ ID NO 82; 1086pp; English.	
PS		
XX	The invention relates to fusion proteins comprising human serum albumin	
CC	(ADH21530) and a therapeutic polypeptide such as a therapeutic protein,	
CC	antibody or peptide or their variants or fragments. The therapeutic	
CC	protein may be fused to the N-terminus, the C-terminus or both termini of	
CC	albumin via a linker. The albumin component of the fusion proteins	
CC	prolongs the shelf-life and the in vitro and vivo biological activity of	
CC	the proteins compared with those of the corresponding therapeutic	
CC	proteins on their own. The invention also relates to nucleic acids	
CC	encoding albumin fusion proteins, vectors and host cells comprising an	
CC	albumin fusion protein nucleic acid, compositions and kits comprising an	
CC	albumin fusion protein, the method of extending the shelf-life of a	
CC	therapeutic protein by fusion with albumin, and the treatment of disease	
CC	using an albumin fusion protein. The albumin fusion proteins may be used	
CC	in the treatment of metabolic/endocrine disorders, diabetes and diabetes-	
CC	related conditions. Specifically the albumin fusion proteins may be used	
CC	to treat type 1 and type 2 diabetes, hyperglycaemia, neural disorders	
CC	(especially neuropathy), retinopathy, cardiovascular disorders	
CC	(especially heart disease, renal disorders and obesity). The proteins may	
CC	also be used in a method of maintaining a basal glucose level in a	
CC	patient and in a method for losing weight. The present sequence is	
CC	related to the invention.	
XX		
SQ	Sequence 876 AA;	
Query Match 83.3%; Score 3152.8; DB 7; Length 876;		
Best Local Similarity 70.4%; Pred. No. 7.9e-73;		
Matches 622; Conservative 13; Mismatches 36; Indels 212; Gaps 10;		
Qy	1 DAHKSEVAHRPKDGENFKALVLIAPYALVLIQPCPFEDHVKLVNEVTEFAKTCVADSEAE 60	
Db	25 DAHKSEVAHRPKDGENFKALVLIAPYALVLIQPCPFEDHVKLVNEVTEFAKTCVADSEAE 84	
Qy	61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPNLRLVRPEV 120	
Db	85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPNLRLVRPEV 144	
Qy	121 DVMTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLLP 180	
Db	145 DVMTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLLP 204	
Qy	181 KLDELREGKASSAKQRIKCSLQKFGERAFAKAVARLSQRFPAKFAEVSKLVTLTK 240	
Db	205 KLDELREGKASSAKQRIKCSLQKFGERAFAKAVARLSQRFPAKFAEVSKLVTLTK 264	
Qy	241 VHTCCGDLLECADDRADLAKYICENQDSISSKLKCCCKPILKSHCTAEVNDMPA 300	
Db	265 VHTCCGDLLECADDRADLAKYICENQDSISSKLKCCCKPILKSHCTAEVNDMPA 324	
Qy	301 DLPSLAADFVSKDVKCYAEAKDVFGLGMFLYFYARRHPDYSVVLLRLAKTYETTLK 360	
Db	325 DLPSLAADFVSKDVKCYAEAKDVFGLGMFLYFYARRHPDYSVVLLRLAKTYETTLK 384	
Qy	361 CAADPHECYAKVDFEFLVVEEONLIKQNCLEFEQLGEYKFNALLVRYTKVQPVST 420	
Db	385 CAADPHECYAKVDFEFLVVEEONLIKQNCLEFEQLGEYKFNALLVRYTKVQPVST 444	
Qy	421 PTLVEVSNLGVSKCKCHPEAKMPCAEADYLSVNLQCLVLEKTPVSDRVTKCCTES 480	
Db	445 PTLVEVSNLGVSKCKCHPEAKMPCAEADYLSVNLQCLVLEKTPVSDRVTKCCTES 504	
Qy	481 LVNRRPCFSALEVDVTFYVKEFNAETFTFHADICTLSEKERQIKKQTAALVELVKHKPKAT 540	

Db	505 LVNRRPCFSALEVDVTFYVKEFNAETFTFHADICTLSEKERQIKKQTAALVELVKHKPKAT 564	
Qy	541 KEQLKAVMDDDFAAFVEKCKCKADDKETCFABEGKLVAAQAALGLMSPLREV----- 592	
Db	565 KEQLKAVMDDDFAAFVEKCKCKADDKETCFABEGKLVAAQAALGL-----LEVAETTPW 620	
Qy	593 -----PCSHALPOGLSPGQVLIIVRGLVLOEPKHFT----- 622	
Db	621 RDAETGERLVCAQCPGPTFVQRPERRDSTTCGP-----CPRHYTQFMYLER 669	
Qy	623 -----VSLRDOAAHA----- 632	
Db	670 CRYCNVLCGEREEARACHATHNRCACRTGFFAHAGFCLEHASCPPGAGVIAPGTPSQN 729	
Qy	633 -----PVTLRAS----- 639	
Db	730 TQCPQCPGPTFFSASSSSSSSQCPHNRCTALGLALNVPGSSSHDTLCTCTGPFPLSTRVPG 769	
Qy	640 -----FADRTLQWISR-----WG-----OKKLIISAPFLFYPO 666	
Db	790 ABECERAVIDFVAFODISIKRLQRLQLEAPGEGWGPTPRAGRAALQLKL-----R 840	
Qy	667 RFEVILLFQEGGLKALANGQGLGATSMNQQALEQLRELRIISG 709	
Db	841 RRLTELLGAGDQGALLVRL-----LQALRVARMPG 869	
RESULT 6		
ADF16517	ID ADF16517 standard; protein; 819 AA.	
XX	AC ADF16517;	
DT	12-FEB-2004 (first entry)	
XX	Human albumin therapeutic fusion protein SeqID1614.	
DE	albumin fusion protein; albumin activity; human serum albumin;	
KW	serum osmotic pressure; shelf-life; stability; antidiabetic;	
KW	gene therapy; diabetes mellitus; human.	
OS	Chimeric.	
OS	Homo sapiens.	
XX	WO2003060071-A2.	
PN	24-JUL-2003.	
XX	23-DEC-2002; 2002WO-US040891.	
XX	21-DEC-2001; 2001US-0341811P.	
PR	24-JAN-2002; 2002US-0350358P.	
PR	28-JAN-2002; 2002US-0351360P.	
PR	26-FEB-2002; 2002US-0359370P.	
PR	28-FEB-2002; 2002US-0360000P.	
PR	27-MAR-2002; 2002US-0367500P.	
PR	08-APR-2002; 2002US-0370227P.	
PR	10-MAY-2002; 2002US-0378950P.	
PR	24-MAY-2002; 2002US-0382617P.	
PR	05-JUN-2002; 2002US-0383123P.	
PR	10-JUL-2002; 2002US-0394625P.	
PR	24-JUL-2002; 2002US-0398008P.	
PR	09-AUG-2002; 2002US-0402131P.	
PR	13-AUG-2002; 2002US-0411355P.	
PR	18-SEP-2002; 2002US-0411426P.	
PR	02-OCT-2002; 2002US-0414984P.	
PR	11-OCT-2002; 2002US-0417611P.	
PR	23-OCT-2002; 2002US-0420246P.	
PR	05-NOV-2002; 2002US-0423623P.	
XX	(HUMA-) HUMAN GENOME SCI INC.	
PA		



PA (DELZ ) DELTA BIOTECHNOLOGY LTD.  
XX (PRIN-) PRINCIPIA PHARM CORP.  
PI Ballance DJ, Turner AJ, Rosen CA, Haseltine WA;  
XX WPI; 2003-598517/56.  
DR New albumin fusion protein, useful for preparing a composition for  
XX treating diabetes mellitus.  
XX Example 4; SEQ ID NO 1614; 24pp; English.  
XX This invention relates to a novel albumin fusion protein having albumin  
CC or biological activity. Human serum albumin is responsible for a  
CC significant proportion of the osmotic pressure of serum and also  
CC functions as a carrier of endogenous and exogenous ligands. The fusion of  
CC albumin to a therapeutic protein may increase shelf-life and stability of  
CC the therapeutic protein. The albumin fusion protein of the invention may  
CC allow production of compositions with antidiabetic activity whilst the  
CC nucleotide sequence which encodes it may be useful for gene therapy. The  
CC albumin fusion protein is useful for preparing a composition for treating  
CC diabetes mellitus. The present sequence is the amino acid sequence of a  
CC novel full-length human albumin therapeutic fusion protein of the  
CC invention. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/publishedpct\_sequences  
XX  
XX Sequence 819 AA;  
Query Match 83.3%; Score 3152.3; DB 7; Length 819;  
Best Local Similarity 76.6%; Pred. NO. 7.2e-73;  
Matches 623; Conservative 16; Mismatches 37; Indels 137; Gaps 12;  
QY 1 DAHSEVAHRFDLGEENFKALVLAFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAG 60  
DB 25 DAHSEVAHRFDLGEENFKALVLAFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAG 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKKDNDPNLRLVRPEV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKKDNDPNLRLVRPEV 144  
QY 121 DVMCTAFHNDNETFLKYLVEIARRHPYFYAPELLFFAKRYKAAPTECCQAADKAACLLP 180  
DB 145 DVMCTAFHNDNETFLKYLVEIARRHPYFYAPELLFFAKRYKAAPTECCQAADKAACLLP 204  
QY 181 KLDELDRSGKASSAKORLKASLOKFGGFAKAWARLSORFPKAEFAEYVKLVTDLTK 240  
DB 205 KLDELDRSGKASSAKORLKASLOKFGGFAKAWARLSORFPKAEFAEYVKLVTDLTK 264  
QY 241 VHTECHGDLLECADDRLAKYICENQDSISSKLECCCEKPLEKSHCIAEVENDENMPA 300  
DB 265 VHTECHGDLLECADDRLAKYICENQDSISSKLECCCEKPLEKSHCIAEVENDENMPA 324  
QY 301 DLPLSLAADPVESKVKYAEAKDVFGLMFLVEYARRHPDYSVLLLLRLAKTYETTLK 360  
DB 325 DLPLSLAADPVESKVKYAEAKDVFGLMFLVEYARRHPDYSVLLLLRLAKTYETTLK 384  
QY 361 CAAADPHCEYAKVDFEFLVEEPONLTKQNCLEFQELGEYKFONALLVRYTKVPQVST 420  
DB 385 CAAADPHCEYAKVDFEFLVEEPONLTKQNCLEFQELGEYKFONALLVRYTKVPQVST 444  
QY 421 PTLVSVRNGLGVSKCKHPKAKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 445 PTLVSVRNGLGVSKCKHPKAKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504  
QY 481 LVNRRPCFSALVDETYVPKEFNAETFTPHADICTLSEKEROIKQTALVELVKHKPKAT 540  
DB 505 LVNRRPCFSALVDETYVPKEFNAETFTPHADICTLSEKEROIKQTALVELVKHKPKAT 564  
QY 541 KQOLKAVMDDDFAAFVEKCKADDKETCFABEGKLVAAASQAALGL-----MSPLREV 592  
DB 565 KQOLKAVMDDDFAAFVEKCKADDKETCFABEGKLVAAASQAALGLATGLTSLDPSLEI 624

QY 593 -----PCSHALPOGL 602  
DB 625 YKQMFVRRREQLLAKNLQALNDIHQQYKILDVMLKGLFKVLEDSRTVLTAADVLPDGP 684  
QY 603 SPQGVIIIRGLVLQEPKHTVSLRQAAHAPVTLIRASFADRTLQ-----W- 647  
DB 685 FP-----QDEK-----LKDAFSHV-VENTAFFGDVLRFPRIHVHYFDHNSNN 727  
QY 648 -ISRWGKKLISAPFLPYQRPFEVLLFOEGGLK--LALNGQGLG----- 690  
DB 728 LLIRWG-----IS-----FCNQTGVNQGHSPILSLMAQBLGISBKDSNFQNPFKI 774  
QY 691 -----ATSMNQAL-----EQLELR 706  
DB 775 DRTEFIPSTDPFQKALREKREKRRKKEKRKEIR 807  
RESULT 7  
ADF16519 ID ADF16519 standard; protein; 819 AA.  
XX AC ADF16519;  
XX 12-FEB-2004 (first entry)  
XX Human albumin therapeutic fusion protein SeqID1616.  
XX albumin fusion protein; albumin activity; human serum albumin;  
XX serum osmotic pressure; shelf-life; stability; antidiabetic;  
XX gene therapy; diabetes mellitus; human.  
XX Chimeric.  
XX Homo sapiens.  
XX WO2003060071-A2.  
XX 24-JUL-2003.  
XX 23-DEC-2002; 2002WO-US040891.  
XX 21-DEC-2001; 2001US-0341811P.  
XX 24-JAN-2002; 2002US-0350358P.  
XX 28-JAN-2002; 2002US-0351360P.  
XX 26-FEB-2002; 2002US-0359370P.  
XX 28-FEB-2002; 2002US-0360000P.  
XX 27-MAR-2002; 2002US-0367500P.  
XX 08-APR-2002; 2002US-0370237P.  
XX 10-MAY-2002; 2002US-0378950P.  
XX 24-MAY-2002; 2002US-0382617P.  
XX 28-MAY-2002; 2002US-0383123P.  
XX 05-JUN-2002; 2002US-0385708P.  
XX 10-JUL-2002; 2002US-0394625P.  
XX 24-JUL-2002; 2002US-0398008P.  
XX 09-AUG-2002; 2002US-0402131P.  
XX 13-AUG-2002; 2002US-0402708P.  
XX 18-SEP-2002; 2002US-0411355P.  
XX 18-SEP-2002; 2002US-0411426P.  
XX 02-OCT-2002; 2002US-0414984P.  
XX 11-OCT-2002; 2002US-0417611P.  
XX 23-OCT-2002; 2002US-0420246P.  
XX 05-NOV-2002; 2002US-0423623P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX (DELZ ) DELTA BIOTECHNOLOGY LTD.  
XX (PRIN-) PRINCIPIA PHARM CORP.  
XX Ballance DJ, Turner AJ, Rosen CA, Haseltine WA;  
XX WPI; 2003-598517/56.  
XX New albumin fusion protein, useful for preparing a composition for  
XX treating diabetes mellitus.



CC using an albumin fusion protein. The albumin fusion proteins may be used  
CC in the treatment of metabolic/endocrine disorders, diabetes and diabetes-  
CC related conditions. Specifically the albumin fusion proteins may be used  
CC to treat type 1 and type 2 diabetes, hyperglycaemia, neural disorders  
CC (especially neuropathy), retinopathy, cardiovascular disorders  
CC (especially heart disease), renal disorders and obesity. The proteins may  
CC also be used in a method of maintaining a basal glucose level in a  
CC patient and in a method for losing weight. The present sequence is  
CC related to the invention.

XX SQ Sequence 819 AA;

Query Match 83.3%; Score 3152.3; DB 7; Length 819;  
Best Local Similarity 76.6%; Pred. No. 7.2e-73;  
Matches 623; Conservative 16; Mismatches 37; Indels 137; Gaps 12;

Qy 1 DAHSEVAHREKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADSEAE 60  
Db 25 DAHSEVAHREKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADSEAE 84  
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNPMLPRLVRPEV 120  
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNPMLPRLVRPEV 144  
Qy 121 DVMTAFHDNEETFLKKLYEIAIRRHYPFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
Db 145 DVMTAFHDNEETFLKKLYEIAIRRHYPFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 204  
Qy 181 KLDELDEGKASSAKQRIKCSLQKFGGERAPKAWAVARLSORFPKAEFAEYSKLVTDITK 240  
Db 205 KLDELDEGKASSAKQRIKCSLQKFGGERAPKAWAVARLSORFPKAEFAEYSKLVTDITK 264  
Qy 241 VHTCECHDLECCADRDADLAKYICENODSISLKECCERPLEKSHCIAEVENDEMPA 300  
Db 265 VHTCECHDLECCADRDADLAKYICENODSISLKECCERPLEKSHCIAEVENDEMPA 324  
Qy 301 DLPSLAADFVESKDVCKNYAEAKVFLGMFLYEYARRHPDYSVLLLLRLAKTYETTLK 360  
Db 325 DLPSLAADFVESKDVCKNYAEAKVFLGMFLYEYARRHPDYSVLLLLRLAKTYETTLK 384  
Qy 361 CAADPHCYAKVDFEPLVEEPPONLIKQNCLEFQOLGEYKFNALLVRYTKKVPQVST 420  
Db 385 CAADPHCYAKVDFEPLVEEPPONLIKQNCLEFQOLGEYKFNALLVRYTKKVPQVST 444  
Qy 421 PTLVEVSRLNGKVGSKCKHPKAKMPCAEEDVLSVLNQLCVLHEKTPVSDRVTKCCTES 480  
Db 445 PTLVEVSRLNGKVGSKCKHPKAKMPCAEEDVLSVLNQLCVLHEKTPVSDRVTKCCTES 504  
Qy 481 LVNRRPCFSALEVDYVYPKEFNAETFTFHADICTLSEKERQIKQTALVELVKHKPKAT 540  
Db 505 LVNRRPCFSALEVDYVYPKEFNAETFTFHADICTLSEKERQIKQTALVELVKHKPKAT 564  
Qy 541 KEQLKAVNDDFAAFVEKCKADDKTCFAEBGKKLVAASQAALGI-----NSPRLV 592  
Db 565 KEQLKAVNDDFAAFVEKCKADDKTCFAEBGKKLVAASQAALGIATGLRTSLDPSLEI 624  
Qy 593 -----QDEK-----LKDASHV-VENTAFGDVVRFPRIYVHYFDHNSWN 727  
Db 625 YKMFVEVRREQLLAKNLALQNDTHQYKILDMVLMKGLFKVLEDSRVTILTAADVLPDGP 684  
Qy 603 SPQGVIIVRGLVQLPQKFTVSLRQAAHAPVTLRASPADRTLQ-----W- 647  
Db 685 PP-----QDEK-----LKDASHV-VENTAFGDVVRFPRIYVHYFDHNSWN 727  
Qy 648 -ISRNGQKLLISAPFLFPQRFVEVLLLPQEGGLK--LALNGQGLG----- 690  
Db 728 LLIRNG-----IS-----FCNQTGVFNQGPSILSLMAQELGISEKDSNFQNPFKI 774  
Qy 691 -----ATSMNQAL-----EQRLER 706  
Db 775 DRTEFIPSTDPPQKALREBEKKRKKKEIR 807

RESULT 9  
ID ADH21808 standard; protein; 819 AA.  
AC ADH21808;  
XX  
DT 11-MAR-2004 (first entry)  
XX  
DE Human albumin/HCE1P80 fragment fusion protein, SEQ ID NO:605.  
XX  
KW Fusion protein; human serum albumin; HSA; therapeutic protein;  
KW shelf-life; in vitro biological activity; in vivo biological activity;  
KW metabolic disorder; endocrine disorder; diabetes; type 1; type 2;  
KW diabetes-related condition; hyperglycaemia; neural disorder; neuropathy;  
KW retinopathy; cardiovascular disorder; heart disease; renal disorder;  
KW obesity; glucose level maintenance; weight loss; antidiabetic; cardiatic;  
KW anorectic; ophthalmological; gene therapy.  
XX  
OS Chimeric.  
OS Homo sapiens.  
XX  
FN WO2003059934-A2.  
XX  
XX 24-JUL-2003.  
PD  
XX 23-DEC-2002; 2002WO-US040892.  
XX  
XX 21-DEC-2001; 2001US-0341811P.  
PR  
XX 24-JAN-2002; 2002US-0350358P.  
PR  
XX 26-FEB-2002; 2002US-0359370P.  
PR  
XX 28-FEB-2002; 2002US-0360000P.  
PR  
XX 27-MAR-2002; 2002US-0367500P.  
PR  
XX 08-APR-2002; 2002US-0370227P.  
PR  
XX 10-MAY-2002; 2002US-0378950P.  
PR  
XX 24-JUL-2002; 2002US-0398008P.  
PR  
XX 09-AUG-2002; 2002US-0402131P.  
PR  
XX 13-AUG-2002; 2002US-0402708P.  
PR  
XX 18-SEP-2002; 2002US-0411355P.  
PR  
XX 02-OCT-2002; 2002US-0414984P.  
PR  
XX 11-OCT-2002; 2002US-0417611P.  
PR  
XX 23-OCT-2002; 2002US-0420246P.  
PR  
XX 05-NOV-2002; 2002US-0423623P.  
XX  
(HUMA-) HUMAN GENOME SCI INC.  
PA  
XX Rosen CA, Haseltine WA;  
PI  
XX WPI; 2003-598501/56.  
DR  
XX New albumin fusion protein, useful for preparing a composition for  
PT treating diabetes mellitus.  
PT  
XX Disclosure; SEQ ID NO 605; 1086pp; English.  
PS  
XX The invention relates to fusion proteins comprising human serum albumin  
CC (ADH21530) and a therapeutic polypeptide such as a therapeutic protein,  
CC antibody or peptide or their variants or fragments. The therapeutic  
CC protein may be fused to the N-terminus, the C-terminus or both termini  
CC of albumin via a linker. The albumin component of the fusion proteins  
CC prolongs the shelf-life and the in vitro and vivo biological activity of  
CC the proteins compared with those of the corresponding therapeutic  
CC proteins on their own. The invention also relates to nucleic acids  
CC encoding albumin fusion proteins, vectors and host cells comprising an  
CC albumin fusion protein nucleic acid, compositions and kits comprising an  
CC albumin fusion protein, the method of extending the shelf-life of a  
CC therapeutic protein by fusion with albumin, and the treatment of disease  
CC using an albumin fusion protein. The albumin fusion proteins may be used  
CC in the treatment of metabolic/endocrine disorders, diabetes and diabetes-  
CC related conditions. Specifically the albumin fusion proteins may be used  
CC to treat type 1 and type 2 diabetes, hyperglycaemia, neural disorders  
CC (especially neuropathy), retinopathy, cardiovascular disorders  
CC (especially heart disease, renal disorders and obesity. The proteins may  
CC also be used in a method of maintaining a basal glucose level in a

CC	patient and in a method for losing weight. The present sequence is	
CC	related to the invention.	
XX		
QQ	Sequence 819 AA;	
	Query Match 83.3%; Score 3152.3; DB 7; Length 819;	
	Best Local Similarity 76.3%; Pred. No. 7.2e-73;	
	Matches 623; Conservative 16; Mismatches 37; Indels 137; Gaps 12;	
QY	1 DAHSEVAHFKDLGEENFKALVLIAPAFQYLCQCPPEFHVKLVNEVTEFAKTCVADESAAE 60	XX
DB	25 DAHSEVAHFKDLGEENFKALVLIAPAFQYLCQCPPEFHVKLVNEVTEFAKTCVADESAAE 84	DE
QY	61 NCDKSLHTLFGDKLCTVATIRYTGEMADCAKQEPERNECFLOHODDNNPLRLVRPEV 120	XX
DB	85 NCDKSLHTLFGDKLCTVATIRYTGEMADCAKQEPERNECFLOHODDNNPLRLVRPEV 144	XX
QY	121 DVMTAFHDNEETFLKKLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLLP 180	XX
DB	145 DVMTAFHDNEETFLKKLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLLP 204	XX
QY	181 KLDELDEGKASSAKQRLKASQKQGERAFKAWAVARLSQRPFAEVSFKLVTDLTG 240	XX
DB	205 KLDELDEGKASSAKQRLKASQKQGERAFKAWAVARLSQRPFAEVSFKLVTDLTG 264	XX
QY	241 VHTCCGHDLLCADDRAADLAKYICENQDISISSKLKCCCKPFLLEKSHCIAEVENDEMPA 300	XX
DB	265 VHTCCGHDLLCADDRAADLAKYICENQDISISSKLKCCCKPFLLEKSHCIAEVENDEMPA 324	XX
QY	301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYEVARRHPDYSVVLLRLAKTETTTLEK 360	XX
DB	325 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYEVARRHPDYSVVLLRLAKTETTTLEK 384	XX
QY	361 CAAADPHCEYAKVDFEKPVEEPQNLIKONCELFEOQLGEYKFNQALLVRYTKVPQVST 420	XX
DB	385 CAAADPHCEYAKVDFEKPVEEPQNLIKONCELFEOQLGEYKFNQALLVRYTKVPQVST 444	XX
QY	421 PTLVEVSRNLGKVGSKCKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCTES 480	XX
DB	445 PTLVEVSRNLGKVGSKCKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCTES 504	XX
QY	481 LVNRRPCFSALEVDVETVVPKEFNAETFTFHADICTLSEKERQIKQOTALVELVKHKKPAT 540	XX
DB	505 LVNRRPCFSALEVDVETVVPKEFNAETFTFHADICTLSEKERQIKQOTALVELVKHKKPAT 564	XX
QY	541 KEQLKAYMDDFAAFVEKCCCKADKCTCFABEGKCLVAASQAALGL-----MSRPLEV 592	XX
DB	565 KEQLKAYMDDFAAFVEKCCCKADKCTCFABEGKCLVAASQAALGLATGTLRTSLDPSLEI 624	XX
QY	593 -----PCSHALPQGL 602	XX
DB	625 YKMFVEVRREQLLAKNLAKLNDIHOQYKILDVMLKGLFKVLBDSRTVLTAADVLDPGP 684	XX
QY	603 SPQGVIIVRGVLQEPKHFTVSLRDQAHAHAPVTLRASFADETLQ-----W- 647	XX
DB	685 FP-----ODEK-----LKDAFSHV-VENTAFPGDVLLRPPIVHYVFDHNSWN 727	XX
QY	648 -ISRWGOKLISAPFLYFPQRFVFLLLFOEGGLK--LALNGQGLG----- 690	XX
DB	728 LLIRWG-----IS-----FCNQTVFNQGHSPILSLMAQELGLISEKDSNPFQNFKI 774	XX
QY	691 -----ATSMNQAL-----RQLREL 706	XX
DB	775 DRTEFIPSTPQKALREEEKRKKKEIR 807	XX
	RESULT 10	
AD	ADF16560	
ID	ADF16560 standard; protein; 794 AA.	
XX		
AC	ADF16560;	
XX		
DT	12-FEB-2004 (first entry)	

Human albumin therapeutic fusion protein SeqID1657.

albumin fusion protein; albumin activity; human serum albumin; serum osmotic pressure; shelf-life; stability; antidiabetic; gene therapy; diabetes mellitus; human.

Chimeric.

Homo sapiens.

WO2003060071-A2.

24-JUL-2003.

23-DEC-2002; 2002WO-US040891.

21-DEC-2001; 2001US-0341811P.

24-JAN-2002; 2002US-0350358P.

28-JAN-2002; 2002US-0351360P.

28-FEB-2002; 2002US-0359370P.

28-FEB-2002; 2002US-0360000P.

27-MAR-2002; 2002US-0367500P.

08-APR-2002; 2002US-0370227P.

10-MAY-2002; 2002US-0378950P.

24-MAY-2002; 2002US-0382617P.

28-MAY-2002; 2002US-0383123P.

05-JUN-2002; 2002US-0385708P.

10-JUL-2002; 2002US-0394625P.

24-JUL-2002; 2002US-0398008P.

09-AUG-2002; 2002US-0402131P.

13-AUG-2002; 2002US-0402708P.

18-SEP-2002; 2002US-0411355P.

18-SEP-2002; 2002US-0411426P.

02-OCT-2002; 2002US-0414984P.

11-OCT-2002; 2002US-0417611P.

23-OCT-2002; 2002US-0420246P.

05-NOV-2002; 2002US-0423623P.

(HUMA-) HUMAN GENOME SCI INC.  
(DELZ ) DELTA BIOTECHNOLOGY LTD.  
(PRIN-) PRINCIPIA PHARM CORP.

Ballance DJ, Turner AJ, Rosen CA, Haseltine WA;  
WFI; 2003-598517/56.

New albumin fusion protein, useful for preparing a composition for treating diabetes mellitus.

Example 4; SEQ ID NO 1657; 24pp; English.

This invention relates to a novel albumin fusion protein having albumin or biological activity. Human serum albumin is responsible for a significant proportion of the osmotic pressure of serum and also functions as a carrier of endogenous and exogenous ligands. The fusion of albumin to a therapeutic protein may increase shelf-life and stability of the therapeutic protein. The albumin fusion protein of the invention may allow production of compositions with antidiabetic activity whilst the nucleotide sequence which encodes it may be useful for gene therapy. The albumin fusion protein is useful for preparing a composition for treating diabetes mellitus. The present sequence is the amino acid sequence of a novel full-length human albumin therapeutic fusion protein of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/publishedpct\_sequences

Sequence 794 AA;

Query Match 83.3%; Score 3152.2; DB 7; Length 794;  
Best Local Similarity 76.3%; Pred. No. 6.9e-73;  
Matches 617; Conservative 18; Mismatches 36; Indels 138; Gaps 9;

1 DAHSEVAHFKDLGEENFKALVLIAPAFQYLCQCPPEFHVKLVNEVTEFAKTCVADESAAE 60

Db 25 DAHKSEVAHRFKDGLGEENFKALVLIAPAYLQQCPFFEDHVKLVNEVTEFAKTCVADESAAE 84  
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144  
Qy 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
Db 145 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 204  
Qy 181 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRPFKABFAEVSUVTDLTK 240  
Db 205 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRPFKABFAEVSUVTDLTK 264  
Qy 241 VHTCECHGDLLECCADDDRADLAKYICENQDSSSSKLKCECKEPLLEKSHCIAEVENDEWPA 300  
Db 265 VHTCECHGDLLECCADDDRADLAKYICENQDSSSSKLKCECKEPLLEKSHCIAEVENDEWPA 324  
Qy 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYIYARRHPDYSVVLLRLAKTYETTLK 360  
Db 325 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYIYARRHPDYSVVLLRLAKTYETTLK 384  
Qy 361 CAAADPHECYAKVFEKPLVEEPQNLKQNCLEFQOLGEYKFNALLVRYTKVPQVST 420  
Db 385 CAAADPHECYAKVFEKPLVEEPQNLKQNCLEFQOLGEYKFNALLVRYTKVPQVST 444  
Qy 421 PTLVEVSRNLGKVGSKCKHPKAPKMPCAEDYLSVNLNOLCVLHEKTPVSDRVTKCCTES 480  
Db 445 PTLVEVSRNLGKVGSKCKHPKAPKMPCAEDYLSVNLNOLCVLHEKTPVSDRVTKCCTES 504  
Qy 481 LVNRRPCFSALEVDYVVPKFNATFTFHADICTLSEKERQIKKQTLALVELVKHKKPAT 540  
Db 505 LVNRRPCFSALEVDYVVPKFNATFTFHADICTLSEKERQIKKQTLALVELVKHKKPAT 564  
Qy 541 KEQLKAVMDDDFAFVEKCKCKADDETCFAEKGKLVAAASQAALGLM-----SPRLEV 592  
Db 565 KEQLKAVMDDDFAFVEKCKCKADDETCFAEKGKLVAAASQAALGLMAFTEHSPLTPHRRD 624  
Qy 593 PCSHAL-----PQGLSPG----- 605  
Db 625 LCSRSIWLARKIRSDLTALTALSYVKGHGLNKNINLDSADGMPVASTDQWSELTEABERLQE 684  
Qy 606 -----QVLIIVGLVLQEPKHFTVSLRD--QAHAAPVTLRASFA----- 641  
Db 685 NLQAYRTHVLJAR--LEEDQVHFTPTBEGDFHQAHTLLOVAFAFYQIBELMILLEYKI 743  
Qy 642 -----DRTLQWISRWGKKLISAPFLFYQRPFFVLLLFQEGGLKIA 683  
Db 744 PRNEADGMPINVGDLPEKKL-----WGLKVL-----QE----- 773  
Qy 684 LMGQGLGATSNQQAQLERLIRISGSVQ 712  
Db 774 -----LSQWTVRSIHDLRFTSSHQ 792

RESULT 11

ID ADF16451 standard; protein; 809 AA.

AC ADF16451;

XX 12-FEB-2004 (first entry)

DE Human albumin therapeutic fusion protein SeqID1548.

XX albumin fusion protein; albumin activity; human serum albumin;  
KW serum osmotic pressure; shelf-life; stability; antidiabetic;  
KW gene therapy; diabetes mellitus; human.

OS Chimeric.  
XX Homo sapiens.

PN WO2003060071-A2.  
XX 24-JUL-2003.  
XX 23-DEC-2002; 2002WO-US040891.  
XX 21-DEC-2001; 2001US-0341811P.  
PR 24-JAN-2002; 2002US-0350358P.  
PR 26-JAN-2002; 2002US-0351360P.  
PR 26-FEB-2002; 2002US-0359370P.  
PR 28-FEB-2002; 2002US-0360000P.  
PR 27-MAR-2002; 2002US-0367500P.  
PR 08-APR-2002; 2002US-0370227P.  
PR 10-MAY-2002; 2002US-0378950P.  
PR 24-MAY-2002; 2002US-0382617P.  
PR 28-MAY-2002; 2002US-0383123P.  
PR 05-JUN-2002; 2002US-0385708P.  
PR 10-JUL-2002; 2002US-0394625P.  
PR 24-JUL-2002; 2002US-0398008P.  
PR 09-AUG-2002; 2002US-0402131P.  
PR 13-AUG-2002; 2002US-0402708P.  
PR 18-SEP-2002; 2002US-0411355P.  
PR 18-SEP-2002; 2002US-0411426P.  
PR 02-OCT-2002; 2002US-0414984P.  
PR 11-OCT-2002; 2002US-0417611P.  
PR 23-OCT-2002; 2002US-0420246P.  
PR 05-NOV-2002; 2002US-0423623P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA (DELZ ) DELTA BIOTECHNOLOGY LTD.  
PA (PRIN-) PRINCIPIA PHARM CORP.  
XX Ballance DJ, Turner AJ, Rosen CA, Haseltine WA;  
WPI; 2003-598517/56.  
XX New albumin fusion protein, useful for preparing a composition for  
treating diabetes mellitus.  
XX Example 4; SEQ ID NO 1548; 24pp; English.  
XX This invention relates to a novel albumin fusion protein having albumin  
or biological activity. Human serum albumin is responsible for a  
significant proportion of the osmotic pressure of serum and also  
functions as a carrier of endogenous and exogenous ligands. The fusion of  
albumin to a therapeutic protein may increase shelf-life and stability of  
the therapeutic protein. The albumin fusion protein of the invention may  
allow production of compositions with antidiabetic activity whilst the  
nucleotide sequence which encodes it may be useful for gene therapy. The  
albumin fusion protein is useful for preparing a composition for treating  
diabetes mellitus. The present sequence is the amino acid sequence of a  
novel full-length human albumin therapeutic fusion protein of the  
invention. Note: The sequence data for this patent did not form part of  
the printed specification, but was obtained in electronic format directly  
from WIPO at ftp.wipo.int/pub/publishedpct\_sequences

XX Sequence 809 AA;

Query Match 83.3%; Score 3152.2; DB 7; Length 809;  
Best Local Similarity 76.3%; Pred. No. 7.1e-73;  
Matches 617; Conservative 18; Mismatches 36; Indels 138; Gaps 9;

Qy 1 DAHKSEVAHRFKDGLGEENFKALVLIAPAYLQQCPFFEDHVKLVNEVTEFAKTCVADESAAE 60  
Db 25 DAHKSEVAHRFKDGLGEENFKALVLIAPAYLQQCPFFEDHVKLVNEVTEFAKTCVADESAAE 84  
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144  
Qy 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
Db 145 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 204

QY 181 KLDELDEGKASSAKORLKCASLOKGERAFKAWAVARLSQRPKFAEFAEVSCLVTDLT 240  
Db 205 KLDELDEGKASSAKORLKCASLOKGERAFKAWAVARLSQRPKFAEFAEVSCLVTDLT 264  
QY 241 VHTECHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLEKSHCHICAEVNDMPA 300  
Db 265 VHTECHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLEKSHCHICAEVNDMPA 324  
QY 301 DLPSLAADFVESKDVCKNVAEAKDVFGLMFLYIARRHPDYSVLLRLRAKTYETTLK 360  
Db 325 DLPSLAADFVESKDVCKNVAEAKDVFGLMFLYIARRHPDYSVLLRLRAKTYETTLK 384  
QY 361 CAAADPHECVAKVDFEFPKLVPEPQNI.KONCELFEQLGKYEKFQNALLVRYTKKVPQVST 420  
Db 385 CAAADPHECVAKVDFEFPKLVPEPQNI.KONCELFEQLGKYEKFQNALLVRYTKKVPQVST 444  
QY 421 PTLVEYSRNLKGVSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
Db 445 PTLVEYSRNLKGVSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504  
QY 481 LVNRRPCFSALEVDYTPKFEFNAETFTFHADICTLSEKERQIKKOTALVELVGHKPKAT 540  
Db 505 LVNRRPCFSALEVDYTPKFEFNAETFTFHADICTLSEKERQIKKOTALVELVGHKPKAT 564  
QY 541 KEOLKAVMDDFAAVFEKCKKADDKETCFPAEEGKKLVAASQAALGLM-----SPRLEV 592  
Db 565 KEOLKAVMDDFAAVFEKCKKADDKETCFPAEEGKKLVAASQAALGLMAFTEHSPSLTPHRRD 624  
QY 593 PCSHAL-----PQGLSPG----- 605  
Db 625 LCSRSIWLAKIRSDLTALTESYVHQGLNKNINLDSADGMPVASTDQWSELTEARLQE 684  
QY 606 -----QVIIVGLVLOPKHFTVSLRD--QAHAFTVLRASPA----- 641  
Db 685 NLOAYRTFHVLLAR-LLEDQOVHFTTEGDFHOAIHTLLIQVAFAFYIEELMILLEYKI 743  
QY 642 -----DRTLQWISRWGOKKLISAPFLYPQPFVLLLFQEGGKLA 683  
Db 744 PRNEADGMPINVGGLGFEKKL-----WGLKVL-----OE----- 773  
QY 684 LNQGLGATSMNQALEQKRELKISGSVQ 712  
Db 774 -----LSQWTVRSIHDLRFISHQ 792

RESULT 12

ADFI5070  
ID ADFI5070 standard; protein; 774 AA.  
AC ADFI5070;  
XX  
XX  
DT 12-FEB-2004 (first entry)  
XX Human albumin therapeutic fusion protein SeqID366.  
DE  
XX albumin fusion protein; albumin activity; human serum albumin;  
KW serum osmotic pressure; shelf-life; stability; antidiabetic;  
KW gene therapy; diabetes mellitus; human.  
XX Chimeric.  
OS Homo sapiens.  
XX WO2003060071-A2.  
PN  
XX 24-JUL-2003.  
XX  
XX 23-DEC-2002; 2002WO-US040891.  
XX  
XX 21-DEC-2001; 2001US-0341811P.  
PR 24-JAN-2002; 2002US-0350358P.  
PR 28-JAN-2002; 2002US-0351360P.  
PR 26-FEB-2002; 2002US-0359370P.  
PR

PR 28-FEB-2002; 2002US-0360000P.  
PR 27-MAR-2002; 2002US-0367500P.  
PR 08-APR-2002; 2002US-0370227P.  
PR 10-MAY-2002; 2002US-0378950P.  
PR 24-MAY-2002; 2002US-0384617P.  
PR 28-MAY-2002; 2002US-0383123P.  
PR 05-JUN-2002; 2002US-0385708P.  
PR 10-JUL-2002; 2002US-0394625P.  
PR 24-JUL-2002; 2002US-0398008P.  
PR 09-AUG-2002; 2002US-0402131P.  
PR 13-AUG-2002; 2002US-0402708P.  
PR 18-SEP-2002; 2002US-0411355P.  
PR 18-SEP-2002; 2002US-0411426P.  
PR 02-OCT-2002; 2002US-0414984P.  
PR 11-OCT-2002; 2002US-0417611P.  
PR 23-OCT-2002; 2002US-0420246P.  
PR 05-NOV-2002; 2002US-0423623P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA (DELZ ) DELTA BIOTECHNOLOGY LTD.  
PA (PRIN-) PRINCIPIA PHARM CORP.  
XX  
XX Ballance DJ, Turner AJ, Rosen CA, Haseltine WA;  
PI WPI; 2003-598517/56.  
XX  
XX New albumin fusion protein, useful for preparing a composition for  
PT treating diabetes mellitus.  
XX  
XX Example 4; SEQ ID NO 366; 24pp; English.  
XX  
XX This invention relates to a novel albumin fusion protein having albumin  
CC or biological activity. Human serum albumin is responsible for a  
CC significant proportion of the osmotic pressure of serum and also  
CC functions as a carrier of endogenous and exogenous ligands. The fusion of  
CC albumin to a therapeutic protein may increase shelf-life and stability of  
CC the therapeutic protein. The albumin fusion protein of the invention may  
CC allow production of compositions with antidiabetic activity whilst the  
CC nucleotide sequence which encodes it may be useful for gene therapy. The  
CC albumin fusion protein is useful for preparing a composition for treating  
CC diabetes mellitus. The present sequence is the amino acid sequence of a  
CC novel full-length human albumin therapeutic fusion protein of the  
CC invention. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/publishedpct\_sequences  
XX  
XX Sequence 774 AA;  
SQ  
Query Match 83.3%; Score 3152; DB 7; Length 774;  
Best Local Similarity 80.3%; Pred. No. 6.7e-73;  
Matches 615; Conservative 10; Mismatches 31; Indels 110; Gaps 7;  
QY 1 DAHKSVAHRFKDLGLENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
Db 25 DAHKSVAHRFKDLGLENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144  
QY 121 DVNCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180  
Db 145 DVNCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 204  
QY 181 KLDELDEGKASSAKORLKCASLOKGERAFKAWAVARLSQRPKFAEFAEVSCLVTDLT 240  
Db 205 KLDELDEGKASSAKORLKCASLOKGERAFKAWAVARLSQRPKFAEFAEVSCLVTDLT 264  
QY 241 VHTECHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLEKSHCHICAEVNDMPA 300  
Db 265 VHTECHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLEKSHCHICAEVNDMPA 324  
QY 301 DLPSLAADFVESKDVCKNVAEAKDVFGLMFLYIARRHPDYSVLLRLRAKTYETTLK 360







```
QY 541 KEQLKAVMDDDFAAFVEKCKCKADDETCFAEEGKKLVAAQAALGLMSPRLL----- 590
Db 565 KEQLKAVMDDDFAAFVEKCKCKADDETCFAEEGKKLVAAQAALGLAPPRLLCDSRVLERY 624
QY 591 -----EVPKSHA-----LPQGLS 603
Db 625 LLEAKAEQITTCGAHCSLNEQITVDPDTKVNFFYAKRMEVGGQAVVWQGLLSEAVL 684
QY 604 PGQVIVRGLVLOEP-----KHTVSLRDOQAHAHVTLRASFAADRTLQWISRW 651
Db 685 RGQALLVQSSQPWEPLQLQHVDKAVGLRSLTLLR-----ALRA----- 723
QY 652 GQKKLISAP-----FLFYPPQFFEVLLLFQEGGKL 682
Db 724 -QKEAISPPDPAASAPLRTITADTF-----RKLPVYSNLFURGKLKL 764

RESULT 14
ADF15006
ID ADF15006 standard; protein; 774 AA.
AC ADF15006;
XX
DT 12-FEB-2004 (first entry)
DE Human albumin therapeutic fusion protein SeqID302.
KW albumin fusion protein; albumin activity; human serum albumin;
KW serum osmotic pressure; shelf-life; stability; antidiabetic;
KW gene therapy; diabetes mellitus; human.
XX
OS Chimeric.
OS Homo sapiens.
XX
PN WO2003060071-A2.
XX
PD 24-JUL-2003.
XX
PF 23-DEC-2002; 2002WO-US040891.
XX
PR 21-DEC-2001; 2001US-0341811P.
PR 24-JAN-2002; 2002US-0350358P.
PR 28-JAN-2002; 2002US-0351360P.
PR 26-FEB-2002; 2002US-0359370P.
PR 28-FEB-2002; 2002US-0360000P.
PR 27-MAR-2002; 2002US-0367500P.
PR 08-APR-2002; 2002US-0370227P.
PR 10-MAY-2002; 2002US-0378950P.
PR 24-MAY-2002; 2002US-0382617P.
PR 28-MAY-2002; 2002US-0383123P.
PR 05-JUN-2002; 2002US-0385708P.
PR 10-JUL-2002; 2002US-0394625P.
PR 24-JUL-2002; 2002US-0398008P.
PR 09-AUG-2002; 2002US-0402131P.
PR 13-AUG-2002; 2002US-0402708P.
PR 18-SEP-2002; 2002US-0411355P.
PR 18-SEP-2002; 2002US-0411426P.
PR 02-OCT-2002; 2002US-0414984P.
PR 11-OCT-2002; 2002US-0417611P.
PR 23-OCT-2002; 2002US-0420246P.
PR 05-NOV-2002; 2002US-0423623P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (DELZ) DELTA BIOTECHNOLOGY LTD.
PA (PRIN-) PRINCIPIA PHARM CORP.
XX
PI Ballance DJ, Turner AJ, Rosen CA, Haseltine WA;
XX
DR WPI; 2003-598517/56.
XX
PT New albumin fusion protein, useful for preparing a composition for
PT treating diabetes mellitus.
```

```
XX
PS
XX
CC This invention relates to a novel albumin fusion protein having albumin
CC or biological activity. Human serum albumin is responsible for a
CC significant proportion of the osmotic pressure of serum and also
CC functions as a carrier of endogenous and exogenous ligands. The fusion of
CC albumin to a therapeutic protein may increase shelf-life and stability of
CC the therapeutic protein. The albumin fusion protein of the invention may
CC allow production of compositions with antidiabetic activity whilst the
CC nucleotide sequence which encodes it may be useful for gene therapy. The
CC albumin fusion protein is useful for preparing a composition for treating
CC diabetes mellitus. The present sequence is the amino acid sequence of a
CC novel full-length human albumin therapeutic fusion protein of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/publishedpct_sequences
XX
SQ Sequence 774 AA;
Query Match 83.2%; Score 3151; DB 7; Length 774;
Best Local Similarity 80.3%; Pred. No. 7.1e-73;
Matches 615; Conservative 9; Mismatches 32; Indels 110; Gaps 7;
QY 1 DAHKSEVAHRFKDLGSENFKALVLIIFAQYLOQCPFEDHVKLVNEVTEFAKTCVADESAAE 60
Db 25 DAHKSEVAHRFKDLGSENFKALVLIIFAQYLOQCPFEDHVKLVNEVTEFAKTCVADESAAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOKDDNPNLRLVRPEV 120
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOKDDNPNLRLVRPEV 144
QY 121 DVNCTAFHDNEETFLKKLYEIAARRHPYFYVAPELLFFAKYKAAFTCCCAAKAACLLP 180
Db 145 DVNCTAFHDNEETFLKKLYEIAARRHPYFYVAPELLFFAKYKAAFTCCCAAKAACLLP 204
QY 181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRFPKFAEFAVSKLVDLT 240
Db 205 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRFPKFAEFAVSKLVDLT 264
QY 241 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLECCCKPLEKSHCHIAEVNDEMPA 300
Db 265 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLECCCKPLEKSHCHIAEVNDEMPA 324
QY 301 DLPSLAADFVESKDVCKNVAEAKDVLGMEFLYFARRHPDYSVVLRLRLAKTVETTLK 360
Db 325 DLPSLAADFVESKDVCKNVAEAKDVLGMEFLYFARRHPDYSVVLRLRLAKTVETTLK 384
QY 361 CAAADPHECYAKVDFEFPKPLVEEPQNLIKONCELFQELGEYKFNALLVRYTKVPQVST 420
Db 385 CAAADPHECYAKVDFEFPKPLVEEPQNLIKONCELFQELGEYKFNALLVRYTKVPQVST 444
QY 421 PTLVEVSRLGKVGSKCKKHPKAPKAPCAEDYLSVLNQLCVLHEKTPSDRVTKCTES 480
Db 445 PTLVEVSRLGKVGSKCKKHPKAPKAPCAEDYLSVLNQLCVLHEKTPSDRVTKCTES 504
QY 481 LVNRRPCFSALEVDYETVVPKEFNAETTFHADICTLSEKEROIKQOTALVELVGHKPKAT 540
Db 505 LVNRRPCFSALEVDYETVVPKEFNAETTFHADICTLSEKEROIKQOTALVELVGHKPKAT 564
QY 541 KEQLKAVMDDDFAAFVEKCKCKADDETCFAEEGKKLVAAQAALGLMSPRLL----- 590
Db 565 KEQLKAVMDDDFAAFVEKCKCKADDETCFAEEGKKLVAAQAALGLAPPRLLCDSRVLERY 624
QY 591 -----EVPKSHA-----LPQGLS 603
Db 625 LLEAKAEANITTCGAHCSLNEQITVDPDTKVNFFYAKRMEVGGQAVVWQGLLSEAVL 684
QY 604 PGQVIVRGLVLOEP-----KHTVSLRDOQAHAHVTLRASFAADRTLQWISRW 651
Db 685 RGQALLVQSSQPWEPLQLQHVDKAVGLRSLTLLR-----ALRA----- 723
QY 652 GQKKLISAP-----FLFYPPQFFEVLLLFQEGGKL 682
```

724 -QKEAISPPDAASAAPLRTITADTF-----RKLFRVYSNFLRGKLU 764

RESULT 15

ADFL14959  
ID ADF14959 standard; protein; 774 AA.

AC ADF14959;

DT 12-FEB-2004 (first entry)

DE Human albumin therapeutic fusion protein SeqID255.

OS albumin fusion protein; albumin activity; human serum albumin;  
KW serum osmotic pressure; shelf-life; stability; antidiabetic;  
KW gene therapy; diabetes mellitus; human.

OS Chimeric.

OS Homo sapiens.

PN W02003060071-A2.

XX 24-JUL-2003.

XX 23-DEC-2002; 2002WO-US040891.

XX 21-DEC-2001; 2001US-0341811P.

XX 28-JAN-2002; 2002US-0350358P.

XX 26-FEB-2002; 2002US-0351360P.

XX 26-FEB-2002; 2002US-0359370P.

XX 27-MAR-2002; 2002US-0360000P.

XX 08-APR-2002; 2002US-0367500P.

XX 10-MAY-2002; 2002US-0370227P.

XX 24-MAY-2002; 2002US-0378950P.

XX 28-MAY-2002; 2002US-0382617P.

XX 05-JUN-2002; 2002US-0383123P.

XX 10-JUL-2002; 2002US-0385708P.

XX 24-JUL-2002; 2002US-0394625P.

XX 09-AUG-2002; 2002US-0398008P.

XX 13-AUG-2002; 2002US-0402131P.

XX 18-SEP-2002; 2002US-0411355P.

XX 02-OCT-2002; 2002US-0411426P.

XX 11-OCT-2002; 2002US-0414984P.

XX 23-OCT-2002; 2002US-0417611P.

XX 05-NOV-2002; 2002US-0420246P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX (DELZ) DELTA BIOTECHNOLOGY LTD.

XX (PRIN-) PRINCIPIA PHARM CORP.

XX Ballance DJ, Turner AJ, Rosen CA, Haseltine WA;

XX WPI; 2003-598517/56.

XX New albumin fusion protein, useful for preparing a composition for treating diabetes mellitus.

XX Example 4; SEQ ID NO 255; 24pp; English.

CC invention. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/publishedpct\_sequences

XX Sequence 774 AA;

Query Match 83.2%; Score 3151; DB 7; Length 774;  
Best local Similarity 80.5%; Pred. No. 7.1e-73;  
Matches 617; Conservative 9; Mismatches 30; Indels 110; Gaps 8;

Qy 1 DAHKSEVAHRPKDLGEENFKALVLIAPAOYLQOCPPEDHDHVKLVNEVTEFAKTCVADSEAE 60

Db 25 DAHKSEVAHRPKDLGEENFKALVLIAPAOYLQOCPPEDHDHVKLVNEVTEFAKTCVADSEAE 84

Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLLQHKDDNPNLPLRVPEV 120

Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLLQHKDDNPNLPLRVPEV 144

Qy 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180

Db 145 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 204

Qy 181 KLDELDRDEGRKASSAKQRLKASLQKFGERAFKAWAVARLSORFFKABFAEVSKLVTDLTK 240

Db 205 KLDELDRDEGRKASSAKQRLKASLQKFGERAFKAWAVARLSORFFKABFAEVSKLVTDLTK 264

Qy 241 VHTTECHGDLLECCADDDRADLAKYICENQDSISSKLKECCEKPLEKSHCIAEVENDEMPA 300

Db 265 VHTTECHGDLLECCADDDRADLAKYICENQDSISSKLKECCEKPLEKSHCIAEVENDEMPA 324

Qy 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVLLRLAKTYETTLK 360

Db 325 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVLLRLAKTYETTLK 384

Qy 361 CAADPHCEYAKVDFEKPVLVEEPQNLIKONCELFEOQGEYKFNALLVRYTKVPOVST 420

Db 385 CAADPHCEYAKVDFEKPVLVEEPQNLIKONCELFEOQGEYKFNALLVRYTKVPOVST 444

Qy 421 PTLVEVSRNLGKVGSKCKHPKAEKRMPCAEADYLSVNLNQLCVLHEKTPVSDRVTKCCTES 480

Db 445 PTLVEVSRNLGKVGSKCKHPKAEKRMPCAEADYLSVNLNQLCVLHEKTPVSDRVTKCCTES 504

Qy 481 LVNRRPCFSALVEDTVVPKBFNAETTFTHADICTLSEKEROIKQOTALVELVHKHPKAT 540

Db 505 LVNRRPCFSALVEDTVVPKBFNAETTFTHADICTLSEKEROIKQOTALVELVHKHPKAT 564

Qy 541 KEQLKAVMDDFAAFVEKCKADDKETCFABEGKKLVAASQAALGLMSPR----- 589

Db 565 KEQLKAVMDDFAAFVEKCKADDKETCFABEGKKLVAASQAALGLMSPR----- 624

Qy 590 -LEVPCSHALP-----OGLS----- 603

Db 625 LLEAKEAEAITTGCAEHCSLNEALITVPTDKVNFYAKMEVGGQAVEVWQGLALLSEAVL 684

Qy 604 PGQVIVRGLVLQEP-----KHFTVSRDQAAHAPVTLIRASFADRTLOWISRW 651

Db 685 RGQALLVASSQPWEFLQLHVDKAVSGLRSLTLRL-----ALRA----- 723

Qy 652 GQKLLISAP-----FLFVPORFEEVILLFQSGGLKL 682

Db 724 -QKEAISPPDAASAAPLRTITADTF-----RKLFRVYSNFLRGKLU 764

RESULT 16

ADFL15004  
ID ADF15004 standard; protein; 774 AA.

XX ADF15004;

XX DT 12-FEB-2004 (first entry)

XX Human albumin therapeutic fusion protein SeqID300.

XX

CC This invention relates to a novel albumin fusion protein having albumin  
CC or biological activity. Human serum albumin is responsible for a  
CC significant proportion of the osmotic pressure of serum and also  
CC functions as a carrier of endogenous and exogenous ligands. The fusion of  
CC albumin to a therapeutic protein may increase shelf-life and stability of  
CC the therapeutic protein. The albumin fusion protein of the invention may  
CC allow production of compositions with antidiabetic activity whilst the  
CC nucleotide sequence which encodes it may be useful for gene therapy. The  
CC albumin fusion protein is useful for preparing a composition for treating  
CC diabetes mellitus. The present sequence is the amino acid sequence of a  
CC novel full-length human albumin therapeutic fusion protein of the



24-JAN-2002; 2002US-0350358P.  
28-JAN-2002; 2002US-0351360P.  
26-FEB-2002; 2002US-0359370P.  
28-FEB-2002; 2002US-0360000P.  
27-MAR-2002; 2002US-0367500P.  
08-APR-2002; 2002US-0370227P.  
10-MAY-2002; 2002US-0378950P.  
24-MAY-2002; 2002US-0382617P.  
28-MAY-2002; 2002US-0383123P.  
05-JUN-2002; 2002US-0385708P.  
10-JUL-2002; 2002US-0394625P.  
24-JUL-2002; 2002US-0398008P.  
09-AUG-2002; 2002US-0402131P.  
13-AUG-2002; 2002US-0402708P.  
18-SEP-2002; 2002US-0411355P.  
18-SEP-2002; 2002US-0411426P.  
02-OCT-2002; 2002US-0414984P.  
11-OCT-2002; 2002US-0417611P.  
23-OCT-2002; 2002US-0420246P.  
05-NOV-2002; 2002US-0423623P.  
(HUMA-) HUMAN GENOME SCI INC.  
(DELZ ) DELTA BIOTECHNOLOGY LTD.  
(PRIN-) PRINCIPIA PHARM CORP.  
Ballance DJ, Turner AJ, Rosen CA, Haseltine WA;  
WPI; 2003-598517/56.  
New albumin fusion protein, useful for preparing a composition for  
treating diabetes mellitus.  
Example 4; SEQ ID NO 245; 24pp; English.  
This invention relates to a novel albumin fusion protein having albumin  
or biological activity. Human serum albumin is responsible for a  
significant proportion of the osmotic pressure of serum and also  
functions as a carrier of endogenous and exogenous ligands. The fusion of  
albumin to a therapeutic protein may increase shelf-life and stability of  
the therapeutic protein. The albumin fusion protein of the invention may  
allow production of compositions with antidiabetic activity whilst the  
nucleotide sequence which encodes it may be useful for gene therapy. The  
albumin fusion protein is useful for preparing a composition for treating  
diabetes mellitus. The present sequence is the amino acid sequence of a  
novel full-length human albumin therapeutic fusion protein of the  
invention. Note: The sequence data for this patent did not form part of  
the printed specification, but was obtained in electronic format directly  
from WIPO at ftp.wipo.int/pub/publishedpct\_sequences  
Sequence 774 AA;  
Query Match 83.2%; Score 3151; DB 7; Length 774;  
Best Local Similarity 80.3%; Pred. No. 7.1e-73;  
Matches 615; Conservative 9; Mismatches 32; Indels 110; Gaps 7;  
1 DAHKEVAHRFKDLGEENFKALVLIAPYALQYQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
25 DAHKEVAHRFKDLGEENFKALVLIAPYALQYQCCPFEDHVKLVNEVTEFAKTCVADESAAE 84  
61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQPERNECFLOHKDNDPNLPLRLVRPEV 120  
85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQPERNECFLOHKDNDPNLPLRLVRPEV 144  
121 DVMCTAFHNEFTFLKKLYETARHPYFAPPELLFFAKRYKAAFECCQADKAAACLLP 180  
145 DVMCTAFHNEFTFLKKLYETARHPYFAPPELLFFAKRYKAAFECCQADKAAACLLP 204  
181 KLDELURDEKASSAKQRLKASLQKFGERAFAKAWAVARLSQRPKAEFAEVSCLVTDLT 240  
205 KLDELURDEKASSAKQRLKASLQKFGERAFAKAWAVARLSQRPKAEFAEVSCLVTDLT 264  
241 VHTCCGDLLECCADDRADLAKYICENQDSSSKLKECEKPLELLEKSHCIAEVENDEMPA 300

Db 265 VHTCCGDLLECCADDRADLAKYICENQDSSSKLKECEKPLELLEKSHCIAEVENDEMPA 324  
Qy 301 DLPFLAADFVESKDVCKNAYAEAKDVFLGMFLYFYARRHPDYVVVLLRLAKTYTTLEK 360  
Db 325 DLPFLAADFVESKDVCKNAYAEAKDVFLGMFLYFYARRHPDYVVVLLRLAKTYTTLEK 384  
Qy 361 CAAADPHECYAKVDFEPKPLVESPNLIKONCELFEOLGEYKFNALLVRYTKVPOVST 420  
Db 385 CAAADPHECYAKVDFEPKPLVESPNLIKONCELFEOLGEYKFNALLVRYTKVPOVST 444  
Qy 421 PTLVSVSRNLGKVGSKCKKPEAKRMPCAEADYLSVNLQCLVLEHKTTPVSDRVTKCCTES 480  
Db 445 PTLVSVSRNLGKVGSKCKKPEAKRMPCAEADYLSVNLQCLVLEHKTTPVSDRVTKCCTES 504  
Qy 481 LVNRRPCFSALEVDVETVYPKEFNAETFTFHADICTLSEKERQIKKQATLVELVGHKPKAT 540  
Db 505 LVNRRPCFSALEVDVETVYPKEFNAETFTFHADICTLSEKERQIKKQATLVELVGHKPKAT 564  
Qy 541 KEQLKAVMDDFAAVFEKCKKADDDKTCFABEGKLVAAASQAALGMSPRLL 590  
Db 565 KEQLKAVMDDFAAVFEKCKKADDDKTCFABEGKLVAAASQAALGMSPRLL 624  
Qy 591 -----EVPCSHA-----LPQGLS 603  
Db 625 LLEAKEAENITTCGAHCSLNENITVDTKVNFYAKRMEVGOQAVEVWQGLALLSEAVL 684  
Qy 604 PGQVILVRLVQLQEP-----KHFTVSLRDOAAHAPVTLTRASFADRTLQWISRW 651  
Db 685 RGQALLVNSQPEWFLQLHVDKAVSGLSLTLLR-----ALRA----- 723  
Qy 652 GQKLLISAP-----FLYPQPFVLLLFQEGGLKL 682  
Db 724 -QKEAISPPDAASAAPLRTITADTF-----RKLFRVYNSFLRGKLU 764  
RESULT 18  
ADF14960  
ID ADF14960 standard; protein; 774 AA.  
XX ADF14960;  
AC ADF14960;  
XX 12-FEB-2004 (first entry)  
DT Human albumin therapeutic fusion protein SeqID256.  
DE albumin fusion protein; albumin activity; human serum albumin;  
KW serum osmotic pressure; shelf-life; stability; antidiabetic;  
KW gene therapy; diabetes mellitus; human.  
XX Chimeric.  
OS Homo sapiens.  
XX WO2003060071-A2.  
XX 24-JUL-2003.  
XX 23-DEC-2002; 2002WO-US040891.  
XX 21-DEC-2001; 2001US-0341811P.  
PR 24-JAN-2002; 2002US-0350358P.  
PR 28-JAN-2002; 2002US-0351360P.  
PR 26-FEB-2002; 2002US-0359370P.  
PR 28-FEB-2002; 2002US-0360000P.  
PR 27-MAR-2002; 2002US-0367500P.  
PR 08-APR-2002; 2002US-0370227P.  
PR 10-MAY-2002; 2002US-0378950P.  
PR 24-MAY-2002; 2002US-0382617P.  
PR 28-MAY-2002; 2002US-0383123P.  
PR 05-JUN-2002; 2002US-0385708P.  
PR 10-JUL-2002; 2002US-0394625P.  
PR 24-JUL-2002; 2002US-0398008P.  
PR 09-AUG-2002; 2002US-0402131P.  
PR 13-AUG-2002; 2002US-0402708P.



XX New albumin fusion protein, useful for preparing a composition for  
PT treating diabetes mellitus.  
PS Example 4; SEQ ID NO 219; 24pp; English.  
XX This invention relates to a novel albumin fusion protein having albumin  
CC or biological activity. Human serum albumin is responsible for a  
CC significant proportion of the osmotic pressure of serum and also  
CC functions as a carrier of endogenous and exogenous ligands. The fusion of  
CC albumin to a therapeutic protein may increase shelf-life and stability of  
CC the therapeutic protein. The albumin fusion protein of the invention may  
CC allow production of compositions with antidiabetic activity whilst the  
CC albumin fusion protein is useful for preparing a composition for treating  
CC diabetes mellitus. The present sequence is the amino acid sequence of a  
CC novel full-length human albumin therapeutic fusion protein of the  
CC invention. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/publishedpct\_sequences  
XX Sequence 868 AA;  
SQ

Query Match 83.2%; Score 3151; DB 7; Length 868;  
Best Local Similarity 70.5%; Pred. No. 8.6e-73;  
Matches 621; Conservative 17; Mismatches 33; Indels 210; Gaps 11;  
QY 1 DAHSEVAHREPKDGEENFKALVLIAPQYLQCCPFEDHVKLVNEVTEFAKTCVADSSAE 60  
DB 25 DAHSEVAHREPKDGEENFKALVLIAPQYLQCCPFEDHVKLVNEVTEFAKTCVADSSAE 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOKDDNPNLRLVPEV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOKDDNPNLRLVPEV 144  
QY 121 DVMTAFHNEETFLKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAKACLLP 180  
DB 145 DVMTAFHNEETFLKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAKACLLP 204  
QY 181 KLDELREGKASSAKORLKCSLQKFGPERAKAVARLSORFPAEVSRLVTLTK 240  
DB 205 KLDELREGKASSAKORLKCSLQKFGPERAKAVARLSORFPAEVSRLVTLTK 264  
QY 241 VHTCCHGDLLECADRADLAKYICENQDSISSKLKECEKPLEKSHCIAEVENDEMPA 300  
DB 265 VHTCCHGDLLECADRADLAKYICENQDSISSKLKECEKPLEKSHCIAEVENDEMPA 324  
QY 301 DLPSLAADFVSKOVCKNYABAKOVFLGMFLYEVARRHPDYSVLLLRKATYETTLK 360  
DB 325 DLPSLAADFVSKOVCKNYABAKOVFLGMFLYEVARRHPDYSVLLLRKATYETTLK 384  
QY 361 CAADPHCEYAKVDFEPFLVEEPONLTKONCELPQOLGEYKFNALLAVRYTKVPOVST 420  
DB 385 CAADPHCEYAKVDFEPFLVEEPONLTKONCELPQOLGEYKFNALLAVRYTKVPOVST 444  
QY 421 PTLVEVSNLGVSKCKCHPEAKMPCAEDYLSVNLQCLVLEKTPVSDRVTKCCTES 480  
DB 445 PTLVEVSNLGVSKCKCHPEAKMPCAEDYLSVNLQCLVLEKTPVSDRVTKCCTES 504  
QY 481 LVNRRPCFSALEVDYTPVPEFNAETFTFHADICTLSEKQRIKQKQALVELVHKHKAT 540  
DB 505 LVNRRPCFSALEVDYTPVPEFNAETFTFHADICTLSEKQRIKQKQALVELVHKHKAT 564  
QY 541 KEQLKAVMDFAAFVEKCKADDKETCFAEKGKLVAAASQALGLMSPLRLEVP----- 593  
DB 565 KEQLKAVMDFAAFVEKCKADDKETCFAEKGKLVAAASQALGLVA---ETPTYPWRDA 621  
QY 594 -----CSHALPQGLSPQVIVRGLVLOE-----PKHFT----- 622  
DB 622 ETGERLVCAQCP-----GTFVQPCRRDSPTTCPCPRHYTQFWNYLRCR 669  
QY 623 -----VSLRDOAAHA----- 632

Db 670 YCNVLGEREEARACHATHNRACRCRTGFFAHAGFCLEHASCPPGAGVIAPCTPSQNTQ 729  
QY 633 -----PVTLRAS----- 639  
Db 730 CQPCPCTFSASSSSSESQCPHRNCTALGLALNVPGSSSHDTLCTCTGFPPLSTRVPGA 789  
QY 640 -----FADRTLQWISR-----WG-----OKKLISAPFLFYPORF 668  
Db 790 ECERAVIDFVAFQDISIKRLQRLLOALEAPEGPGTTPRAGRAALQLKL-----RRR 840  
QY 669 FEVLLLFQEGGLKALNGQGLGATSMNQQALEQLRELRLISG 709  
Db 841 LTELHGQDQALLVRL-----LQALRVARMPG 867

## RESULT 20

ADH21275  
ID ADH21275 standard; protein; 868 AA.

AC ADH21275;

XX 11-MAR-2004 (first entry)

XX Human albumin/TF6 fragment fusion protein, SEQ ID NO:72.

XX Fusion protein; human serum albumin; HSA; therapeutic protein;  
KW shelf-life; in vitro biological activity; in vivo biological activity;  
KW metabolic disorder; endocrine disorder; diabetes; type 1; type 2;  
KW diabetes-related condition; hyperglycaemia; neural disorder; neuropathy;  
KW retinopathy; cardiovascular disorder; heart disease; renal disorder;  
KW obesity; glucose level maintenance; weight loss; antidiabetic; cardiant;  
KW anorectic; ophthalmological; gene therapy.

XX Chimeric.

OS Homo sapiens.

XX WO2003059934-A2.

XX 24-JUL-2003.

XX 23-DEC-2002; 2002WO-US040892.

XX 21-DEC-2001; 2001US-0341811P.

XX 24-JAN-2002; 2002US-0350358P.

XX 26-FEB-2002; 2002US-0359370P.

XX 28-FEB-2002; 2002US-0360000P.

XX 27-MAR-2002; 2002US-0367500P.

XX 08-APR-2002; 2002US-0370227P.

XX 10-MAY-2002; 2002US-0378950P.

XX 24-JUL-2002; 2002US-0398008P.

XX 09-AUG-2002; 2002US-0402131P.

XX 13-AUG-2002; 2002US-0402708P.

XX 18-SEP-2002; 2002US-0411355P.

XX 02-OCT-2002; 2002US-0414984P.

XX 11-OCT-2002; 2002US-0417611P.

XX 23-OCT-2002; 2002US-0420246P.

XX 05-NOV-2002; 2002US-0423623P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Haseltine WA;

XX WPI; 2003-598501/56.

XX New albumin fusion protein, useful for preparing a composition for

PT treating diabetes mellitus.

XX Disclosure; SEQ ID NO 72; 1086pp; English.

XX The invention relates to fusion proteins comprising human serum albumin  
CC (ADH21530) and a therapeutic polypeptide such as a therapeutic protein,  
CC antibody or peptide or their variants or fragments, the therapeutic  
CC protein may be fused to the N-terminus, the C-terminus or both termini of

CC albumin via a linker. The albumin component of the fusion proteins  
CC prolongs the shelf-life and the in vitro and vivo biological activity of  
CC the proteins compared with those of the corresponding therapeutic  
CC proteins on their own. The invention also relates to nucleic acids  
CC encoding albumin fusion proteins, vectors and host cells comprising an  
CC albumin fusion protein nucleic acid, compositions and kits comprising an  
CC albumin fusion protein, the method of extending the shelf-life of a  
CC therapeutic protein by fusion with albumin, and the treatment of disease  
CC using an albumin fusion protein. The albumin fusion proteins may be used  
CC in the treatment of metabolic/endocrine disorders, diabetes and diabetes-  
CC related conditions. Specifically the albumin fusion proteins may be used  
CC to treat type 1 and type 2 diabetes, hyperglycaemia, neural disorders  
CC (especially neuropathy), retinopathy, cardiovascular disorders  
CC (especially heart disease, renal disorders and obesity. The proteins may  
CC also be used in a method of maintaining a basal glucose level in a  
CC patient and in a method for losing weight. The present sequence is  
CC related to the invention.  
XX  
SQ Sequence 868 AA;

Query Match 83.2%; Score 3151; DB 7; Length 869;  
Best Local Similarity 70.5%; Pred. No. 8.6e-73;  
Matches 621; Conservative 17; Mismatches 33; Indels 210; Gaps 11;

QY 1 DAHSEVAHFPKDLGKFNKALVLIAPAOVLQCPREDHVKLNVETFAKTCVADESAE 60  
DB 25 DAHSEVAHFPKDLGKFNKALVLIAPAOVLQCPREDHVKLNVETFAKTCVADESAE 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144  
QY 121 DVNCTAFHDNEETFLKKYLVEIARRHPYFYVAPELLFFAKRYKAAFTCCQAADKAACLLP 180  
DB 145 DVNCTAFHDNEETFLKKYLVEIARRHPYFYVAPELLFFAKRYKAAFTCCQAADKAACLLP 204  
QY 181 KLDELDEGKASAKQFLKCSAQKGERAFKAWAVARLSQRPFAEFVSKLVLDLTK 240  
DB 205 KLDELDEGKASAKQFLKCSAQKGERAFKAWAVARLSQRPFAEFVSKLVLDLTK 264  
QY 241 VHTTECHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300  
DB 265 VHTTECHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 324  
QY 301 DLPSLAADFVESKDVCKNVAEAKDVFGLMFLYERARHPDYSVVLRLAKTVEITTEK 360  
DB 325 DLPSLAADFVESKDVCKNVAEAKDVFGLMFLYERARHPDYSVVLRLAKTVEITTEK 384  
QY 361 CAADPHECYAKVDFEKPFLVERPQNLKONCELFQOLGEYKFNALLVRYTKVPQVST 420  
DB 385 CAADPHECYAKVDFEKPFLVERPQNLKONCELFQOLGEYKFNALLVRYTKVPQVST 444  
QY 421 PTLVEVSRLNGKVGSKCKHPEAKRMPKABDYLVSVLNQLCVLHEKTPVSDRVTKCTES 480  
DB 445 PTLVEVSRLNGKVGSKCKHPEAKRMPKABDYLVSVLNQLCVLHEKTPVSDRVTKCTES 504  
QY 481 LVNRRPCFSALEVDVETVPKFEANETTFHADICTLSEKERQIKQATALVELVKHPKAT 540  
DB 505 LVNRRPCFSALEVDVETVPKFEANETTFHADICTLSEKERQIKQATALVELVKHPKAT 564  
QY 541 KEOLKAVMDPFAAFVEKCKKADDKETCFABEGKKLVAASQAALGLVA----- 593  
DB 565 KEOLKAVMDPFAAFVEKCKKADDKETCFABEGKKLVAASQAALGLVA---ETPTYPWRDA 621  
QY 594 -----CSHALPQGLSPGVIIIRGLVLOE-----PKHFT----- 622  
DB 622 ETGERLVCAQCCP-----GTFVQPCRRDSTTCGPPCPPHYTFQWNYLERCR 669  
QY 623 -----VSLRDOAHVA----- 632  
DB 670 YCNVLCGERBEARACHATHNRACRRTGFFAHAGFLEHASCPGAGVIAPGTPSQNTQ 729  
QY 633 -----PVTLRAS----- 639

DB 730 CQPCPGTFSASSSSSEQCPHRCNTALGLALNVPGSSSHDTLTCTGFFPLSTRVPGA 789  
QY 640 -----FADRTLQWISR-----WG-----QKLLISAPFLFVPORF 668  
DB 790 ECERAVIDFVAFODISIKRLQRLQLQALEAPGCGPTFRAGRAALQKL-----RRR 840  
QY 669 FEVLILFQSGGKLKALNGQGLGATSMNQALEOLRELRISSG 709  
DB 841 LTELGAQDALLVRL-----LQALRVARMPG 867

RESULT 21  
ADF14922  
ID ADF14922 standard; protein; 874 AA.  
XX  
AC ADF14922;  
XX  
DT 12-FEB-2004 (first entry)  
XX  
DE Human albumin therapeutic fusion protein SeqID218.  
XX  
KW albumin fusion protein; albumin activity; human serum albumin;  
KW serum osmotic pressure; shelf-life; stability; antidiabetic;  
KW gene therapy; diabetes mellitus; human.  
XX  
OS Chimeric.  
OS Homo sapiens.  
XX  
PN WO2003060071-A2.  
XX  
PD 24-JUL-2003.  
XX  
PF 23-DEC-2002; 2002WO-US040891.  
XX  
PR 21-DEC-2001; 2001US-0341811P.  
PR 24-JAN-2002; 2002US-0350358P.  
PR 28-JAN-2002; 2002US-0351360P.  
PR 28-FEB-2002; 2002US-0359370P.  
PR 28-FEB-2002; 2002US-0360000P.  
PR 27-MAR-2002; 2002US-0367500P.  
PR 08-APR-2002; 2002US-0370227P.  
PR 10-MAY-2002; 2002US-0378950P.  
PR 24-MAY-2002; 2002US-0382617P.  
PR 28-MAY-2002; 2002US-0383123P.  
PR 05-JUN-2002; 2002US-0385708P.  
PR 10-JUL-2002; 2002US-0394625P.  
PR 24-JUL-2002; 2002US-0398008P.  
PR 09-AUG-2002; 2002US-0402131P.  
PR 13-AUG-2002; 2002US-0402708P.  
PR 18-SEP-2002; 2002US-0411355P.  
PR 18-SEP-2002; 2002US-0411426P.  
PR 02-OCT-2002; 2002US-0414984P.  
PR 11-OCT-2002; 2002US-0417611P.  
PR 23-OCT-2002; 2002US-0420246P.  
PR 05-NOV-2002; 2002US-0423623P.  
XX  
(HUMA-) HUMAN GENOME SCI INC.  
PA (DELZ ) DELTA BIOTECHNOLOGY LTD.  
PA (PRIN-) PRINCIPIA PHARM CORP.  
XX  
PI Ballance DJ, Turner AJ, Rosen CA, Haseltine WA;  
XX  
XX WPI; 2003-598517/56.  
DR  
XX  
XX New albumin fusion protein, useful for preparing a composition for  
PT treating diabetes mellitus.  
PT  
XX  
PS Example 4; SEQ ID NO 218; 24pp; English.  
XX  
CC This invention relates to a novel albumin fusion protein having albumin  
CC or biological activity. Human serum albumin is responsible for a  
CC significant proportion of the osmotic pressure of serum and also



CC functions as a carrier of endogenous and exogenous ligands. The fusion of  
CC albumin to a therapeutic protein may increase shelf-life and stability of  
CC the therapeutic protein. The albumin fusion protein of the invention may  
CC allow production of compositions with antidiabetic activity whilst the  
CC nucleotide sequence which encodes it may be useful for gene therapy. The  
CC albumin fusion protein is useful for preparing a composition for treating  
CC diabetes mellitus. The present sequence is the amino acid sequence of a  
CC novel full-length human albumin therapeutic fusion protein of the  
CC invention. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/publishedpct\_sequences  
XX  
SQ Sequence 874 AA;

Query Match 83.2%; Score 3151; DB 7; Length 874;  
Best Local Similarity 70.5%; Pred. No. 8.7e-73;  
Matches 621; Conservative 17; Mismatches 33; Indels 210; Gaps 11;

Qy 1 DAHKEVAHRFKDILGEBNFKALVLIAPAYLQOCPPFEDHVKLVNEVTEFAKTCVADSEAE 60  
Db 25 DAHKEVAHRFKDILGEBNFKALVLIAPAYLQOCPPFEDHVKLVNEVTEFAKTCVADSEAE 84  
Qy 61 NCDKSLHTLFGDKLCTVATLEETGEMADCCAKQEPERNECFLOHKDNNPLRLVRPEV 120  
Db 85 NCDKSLHTLFGDKLCTVATLEETGEMADCCAKQEPERNECFLOHKDNNPLRLVRPEV 144  
Qy 121 DVMCTAFHDNBEETFLKYLVIARRHPYFYAPPELLFAKRYKAAFTCCQAAADKAACLLP 180  
Db 145 DVMCTAFHDNBEETFLKYLVIARRHPYFYAPPELLFAKRYKAAFTCCQAAADKAACLLP 204  
Qy 181 KLDELDEGKASSAKQRLKCSLQKFGERAFAKAWAVARLSQRFPAKFAEVSKLVTDLTK 240  
Db 205 KLDELDEGKASSAKQRLKCSLQKFGERAFAKAWAVARLSQRFPAKFAEVSKLVTDLTK 264  
Qy 241 VHTCCGGDLLECCADRDADIAKYICENQDSISSKLKCECKPLEKSHCIAEVENDEMPA 300  
Db 265 VHTCCGGDLLECCADRDADIAKYICENQDSISSKLKCECKPLEKSHCIAEVENDEMPA 324  
Qy 301 DLPLSADPFVSKDVKYAKADVFLGMFLYEVARRHPDYSVLLRLAKTYETTTLEKC 360  
Db 325 DLPLSADPFVSKDVKYAKADVFLGMFLYEVARRHPDYSVLLRLAKTYETTTLEKC 384  
Qy 361 CAADPHECAKVFDEPKLVEEPQNLKONCELFQELGEYKFNALLVRYTKVPQVST 420  
Db 385 CAADPHECAKVFDEPKLVEEPQNLKONCELFQELGEYKFNALLVRYTKVPQVST 444  
Qy 421 PTLVEVSRLGKSGKCKHPEAKMPCAEADYLSVNLQCLVHLHEKTPVSDRVTKCCTES 480  
Db 445 PTLVEVSRLGKSGKCKHPEAKMPCAEADYLSVNLQCLVHLHEKTPVSDRVTKCCTES 504  
Qy 481 LVNRRPCFSALEVDVYVPKEFNAETFTFHADICTLSEKERQIKQATALVELVKHKPKAT 540  
Db 505 LVNRRPCFSALEVDVYVPKEFNAETFTFHADICTLSEKERQIKQATALVELVKHKPKAT 564  
Qy 541 KEQLKAVNMDFAAFVEKCKKADDETCFAEKGKLVAAASQAALGMLSPRLVPP 593  
Db 565 KEQLKAVNMDFAAFVEKCKKADDETCFAEKGKLVAAASQAALGMLV---ETPTVPRDA 621  
Qy 594 -----CSHALPQGLSPGVIIIVGLVLOE-----PKHFT----- 622  
Db 622 ETGERLVCAQCP-----GTFVQPCRRDSTPTCCPCPPRHRYTQFWNYLRCR 669  
Qy 623 -----VSLRDQAAHA----- 632  
Db 670 YCNVLGGBREBARACHATHNACRCRTGTFPAHAGFCLHASCPCPGAGVIAPGTPSQNTQ 729  
Qy 633 -----PVTLRAS----- 639  
Db 730 QCPCPGTFASSSSEQCQPHNCTALGLALNVPSSSHDITLCTSGTFPLSTRVPGA 789  
Qy 640 -----PADTLOWISR-----WG-----OKKLISAPFLFYPORF 668  
Db 790 ECERAVIDFVAFODISIKRLQRLQALBAPEGWGTPPAGRAALQKL-----RRR 840

Qy 669 FEVILLFQEGGLKALNQGLGATSMNQALQELRELISG 709  
Db 841 LTELGAQDQALLVRL-----LQALRVARMFG 867

RESULT 22  
ADH21274  
ID ADH21274 standard; protein; 874 AA.  
XX AC ADH21274;  
XX DT 11-MAR-2004 (first entry)  
XX DE Human albumin/TR6 fragment fusion protein, SEQ ID NO:71.  
XX KW Fusion protein; human serum albumin; HSA; therapeutic protein;  
XX KW shelf-life; in vitro biological activity; in vivo biological activity;  
XX KW metabolic disorder; endocrine disorder; diabetes; type 1; type 2;  
XX KW diabetes-related condition; hyperglycaemia; neural disorder; neuropathy;  
XX KW retinopathy; cardiovascular disorder; heart disease; renal disorder;  
XX KW obesity; glucose level maintenance; weight loss; antidiabetic; cardiant;  
XX KW anorectic; ophthalmological; gene therapy.  
XX OS Chimeric.  
XX OS Homo sapiens.  
XX PN WO2003059934-A2.  
XX PD 24-JUL-2003.  
XX PF 23-DEC-2002; 2002WO-US040892.  
XX PR 21-DEC-2001; 2001US-0341811P.  
XX PR 24-JAN-2002; 2002US-0350358P.  
XX PR 26-FEB-2002; 2002US-0359370P.  
XX PR 28-FEB-2002; 2002US-0360000P.  
XX PR 27-MAR-2002; 2002US-0367500P.  
XX PR 08-APR-2002; 2002US-0370227P.  
XX PR 10-MAY-2002; 2002US-0378950P.  
XX PR 24-JUL-2002; 2002US-0398008P.  
XX PR 09-AUG-2002; 2002US-0402131P.  
XX PR 13-AUG-2002; 2002US-0402708P.  
XX PR 18-SEP-2002; 2002US-0411355P.  
XX PR 02-OCT-2002; 2002US-0414984P.  
XX PR 11-OCT-2002; 2002US-0417611P.  
XX PR 23-OCT-2002; 2002US-0420246P.  
XX PR 05-NOV-2002; 2002US-0423623P.  
XX PA (HUMA-) HUMAN GENOME SCI INC.  
XX PI Rosen CA, Haseltine WA;  
XX WPI; 2003-598501/56.  
XX DR New albumin fusion protein, useful for preparing a composition for  
XX PT treating diabetes mellitus.  
XX PS Disclosure; SEQ ID NO 71; 1086pp; English.  
XX CC The invention relates to fusion proteins comprising human serum albumin  
XX CC (ADH21530) and a therapeutic polypeptide such as a therapeutic protein,  
XX CC antibody or peptide or their variants or fragments. The therapeutic  
XX CC protein may be fused to the N-terminus, the C-terminus or both termini  
XX CC of albumin via a linker. The albumin component of the fusion proteins  
XX CC prolongs the shelf-life and the in vitro and vivo biological activity of  
XX CC the proteins compared with those of the corresponding therapeutic  
XX CC proteins on their own. The invention also relates to nucleic acids  
XX CC encoding albumin fusion proteins, vectors and host cells comprising an  
XX CC albumin fusion protein nucleic acid, compositions and kits comprising an  
XX CC albumin fusion protein, the method of extending the shelf-life of a  
XX CC therapeutic protein by fusion with albumin, and the treatment of disease  
XX CC using an albumin fusion protein. The albumin fusion proteins may be used

in the treatment of metabolic/endocrine disorders, diabetes and diabetes-related conditions. Specifically the albumin fusion proteins may be used to treat type 1 and type 2 diabetes, hyperglycaemia, neural disorders (especially neuropathy), retinopathy, cardiovascular disorders (especially heart disease, renal disorders and obesity). The proteins may also be used in a method of maintaining a basal glucose level in a patient and in a method for losing weight. The present sequence is related to the invention.

Query Match	83.2%	Score 3151;	DB 7;	Length 874;
Best Local Similarity	70.5%;	Pred. No. 8.7e-73;		
Matches 621; Conservative	17;	Mismatches 33;	Indels 210;	Gaps 11;
QY 1	DAHKSEVAHFFKDLGEENFKALVLI	FAQYLOQC	PFEDHVKLVNEVTEFAKTCVADSSAE	60
DB				
QY 25	DAHKSEVAHFFKDLGEENFKALVLI	FAQYLOQC	PFEDHVKLVNEVTEFAKTCVADSSAE	84
DB				
QY 61	NCDKSLHTLFGDKLCTVATIRETY	GEMADCCAKOEP	PERNECFLOHKDDNPNLPLVRPEV	120
DB				
QY 85	NCDKSLHTLFGDKLCTVATIRETY	GEMADCCAKOEP	PERNECFLOHKDDNPNLPLVRPEV	144
DB				
QY 121	DMVCTAFHDNEEFKLVYELIAR	HPHYFYAPELLFFAKRYKAAFT	CECCQAADKAAACLLP	180
DB				
QY 145	DMVCTAFHDNEEFKLVYELIAR	HPHYFYAPELLFFAKRYKAAFT	CECCQAADKAAACLLP	204
DB				
QY 181	KJDELDEGKASSAKQRLKCA	SLQKFGERAFKAWAVARLS	SQRPFKAEVSKLVTDLTK	240
DB				
QY 205	KJDELDEGKASSAKQRLKCA	SLQKFGERAFKAWAVARLS	SQRPFKAEVSKLVTDLTK	264
DB				
QY 241	VHTECCGHDLLLECADDRAD	LAKYICENQDSISSKLKECCE	KPLLEKSHCHTAEVNDEMPA	300
DB				
QY 265	VHTECCGHDLLLECADDRAD	LAKYICENQDSISSKLKECCE	KPLLEKSHCHTAEVNDEMPA	324
DB				
QY 301	DLPSLAADFVESKDVCKNYAE	AKDVFLGMFLYEYARRHPD	YSVLLLRLLARKATYETTLK	360
DB				
QY 325	DLPSLAADFVESKDVCKNYAE	AKDVFLGMFLYEYARRHPD	YSVLLLRLLARKATYETTLK	384
DB				
QY 361	CAAADPHECYAKVFDEFKPL	VEEPQNLIKONCELFEQ	LGKVFQNALLVRYTKKVPQVST	420
DB				
QY 385	CAAADPHECYAKVFDEFKPL	VEEPQNLIKONCELFEQ	LGKVFQNALLVRYTKKVPQVST	444
DB				
QY 421	PTLVEVSRNLGKVGSKCKK	PEAKBPCABDYL	SVLNQLCVLHEKTPVSDRVTKCCTES	480
DB				
QY 445	PTLVEVSRNLGKVGSKCKK	PEAKBPCABDYL	SVLNQLCVLHEKTPVSDRVTKCCTES	504
DB				
QY 481	LVNRRPCFSALEVNDVETVP	KFEFNAETTFTHADICTL	SEKEROIKKQATVELVHKHKPKAT	540
DB				
QY 505	LVNRRPCFSALEVNDVETVP	KFEFNAETTFTHADICTL	SEKEROIKKQATVELVHKHKPKAT	564
DB				
QY 541	KEQLKAVMDFFAAPVEKCK	CAKDDKTCFABEGKKL	VAASQAALGIMSPRLVEP	593
DB				
QY 565	KEQLKAVMDFFAAPVEKCK	CAKDDKTCFABEGKKL	VAASQAALGLVA---ETPTYWRDA	621
DB				
QY 594	-----CSHALPOGLSPGG	VIIVRGLVLOB-	-----PKHFT-----	622
DB				
QY 622	ETGERLVCAQCP	-----GTFVQPCRRDS	PTTCGCPPPRHITQFWNYLRCR	669
DB				
QY 623	-----VSLRDQAHA-	-----	-----	632
DB				
QY 670	YCNVLGGEREEERACHATH	NRACRRTGFFAHAGFC	LEHASCPPGAGVIAPGTPSQNTQ	729
DB				
QY 633	-----PVTLRAS-	-----	-----	639
DB				
QY 730	CQCPCTGTFASASSSSSQ	CPQPNRCHTALGLALN	PGSSSHDTLCTCTGTFPLSTRVPGAE	789
DB				
QY 640	-----FADRTLOWISR-	-----WG-----	-----QKKLISAPFLFYQRF	668
DB				
QY 790	ECERAVIDFVAFQDISIK	RLQLQLQALEAPG	GWGFTPRAGFAALQLKL-----RRR	840
DB				
QY 669	FEVLLLFQEGGLKUALNQ	GGIGATSMNQQALEQL	RELRISG 709	
DB				

CC the printed specification, but was obtained in electronic format directly  
XX from WIPO at ftp.wipo.int/pub/publishedpct\_sequences  
SQ Sequence 880 AA;

Query Match 83.2%; Score 3151; DB 7; Length 880;  
Best Local Similarity 70.5%; Pred. No. 8.8e-73;  
Matches 621; Conservative 17; Mismatches 33; Indels 210; Gaps 11;  
Qy 1 DAHSEVAHFRKDI GEENFKALVLI AFAQYLQOC PFEDHVKLVNEVTEFAKTCVADESA 60  
Db 25 DAHSEVAHFRKDI GEENFKALVLI AFAQYLQOC PFEDHVKLVNEVTEFAKTCVADESA 84  
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCAKQEPERNECFLOHKDDNPNLRLVRPEV 144  
Qy 121 DVMTAFHDNEBTFKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
Db 145 DVMTAFHDNEBTFKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 204  
Qy 181 KLDLRDEGKASSAKORLKASLOKFGERAFAKAVARLSORFPKAFVSKLVTLDTLTK 240  
Db 205 KLDLRDEGKASSAKORLKASLOKFGERAFAKAVARLSORFPKAFVSKLVTLDTLTK 264  
Qy 241 VHTCCCHGDLLECADDRLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 300  
Db 265 VHTCCCHGDLLECADDRLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 324  
Qy 301 DLPSLAADFVSKDVKCNKAYAEAKDVLGMFLYVEYARRHPDYSVVLLRLAKTYETTLK 360  
Db 325 DLPSLAADFVSKDVKCNKAYAEAKDVLGMFLYVEYARRHPDYSVVLLRLAKTYETTLK 384  
Qy 361 CAADPHCYAKVDFEPKPLVEEONLIKONCELFEOQLGEYKFNALLVRYTKVPQVST 420  
Db 385 CAADPHCYAKVDFEPKPLVEEONLIKONCELFEOQLGEYKFNALLVRYTKVPQVST 444  
Qy 421 PTLVEVSRLGKVGSKCCPKHPEAKMPCAEYDLSVLNQLCVLHEKTPVSDRVTKCCTES 480  
Db 445 PTLVEVSRLGKVGSKCCPKHPEAKMPCAEYDLSVLNQLCVLHEKTPVSDRVTKCCTES 504  
Qy 481 LVNRRPCFSALEVDYVYKPEFNAETFTFHADICTLSEKERQIKQTALVELVKHKPKAT 540  
Db 505 LVNRRPCFSALEVDYVYKPEFNAETFTFHADICTLSEKERQIKQTALVELVKHKPKAT 564  
Qy 541 KEQLKAVNMDFAAFVEKCKADDKETCFABEGKKLVAASQAALGLMSPLRLEVP----- 593  
Db 565 KEQLKAVNMDFAAFVEKCKADDKETCFABEGKKLVAASQAALGLVA---ETPTYPWRDA 621  
Qy 594 -----CSHALPQGLSPQGVIIIRGLVLOE-----PKHFT----- 622  
Db 622 ETGERLVCACQFP-----GTFVQRCRRDSTTCGCPPRHYTQFWNYLERCR 669  
Qy 623 -----VSLRDQAAHA----- 632  
Db 670 YCNVLGCEBEARACHATHNACRCRTGFFFAHAGFCLEHASCPPGAGVIAPGTPSQNTQ 729  
Qy 633 -----PVTLRAS----- 639  
Db 730 QCPCPPTGFSASSSSSEQCPHRNCTALGLALNVPSSSHDTLCTSCGTFPLSTRVPGA 789  
Qy 640 -----PADRTLQWISR-----WG-----OKKLISAPFLFYPORF 668  
Db 790 ECERAVIDFVAFQDISIKRLQRLQALRAPEWGTPPRAGRALQKL-----RRR 840  
Qy 669 FVLLLFQEGGKLKALNGQGLGATSNQOALFQELRLRISG 709  
Db 841 LTELIGAQQDGLLVRL-----LQALVARWMPG 867

RESULT 24  
ADF14983  
ID ADF14983 standard; protein; 880 AA.

XX ADF14983;  
XX DT 12-FEB-2004 (first entry)  
XX DE Human albumin therapeutic fusion protein SeqID279.  
XX KW albumin fusion protein; albumin activity; human serum albumin;  
KW serum osmotic pressure; shelf-life; stability; antidiabetic;  
XX gene therapy; diabetes mellitus; human.  
XX OS Chimeric.  
XX OS Homo sapiens.  
XX PN WO2003060071-A2.  
XX PD 24-JUL-2003.  
XX PF 23-DEC-2002; 2002WO-US040891.  
XX PR 21-DEC-2001; 2001US-0341811P.  
PR 24-JAN-2002; 2002US-0350358P.  
PR 28-JAN-2002; 2002US-0351360P.  
PR 26-FEB-2002; 2002US-0359370P.  
PR 28-FEB-2002; 2002US-0360000P.  
PR 27-MAR-2002; 2002US-0367500P.  
PR 08-APR-2002; 2002US-0370227P.  
PR 10-MAY-2002; 2002US-0378950P.  
PR 24-MAY-2002; 2002US-0382617P.  
PR 28-MAY-2002; 2002US-0383123P.  
PR 05-JUN-2002; 2002US-0385708P.  
PR 10-JUL-2002; 2002US-0394625P.  
PR 24-JUL-2002; 2002US-0398008P.  
PR 09-AUG-2002; 2002US-0402131P.  
PR 13-AUG-2002; 2002US-0402708P.  
PR 18-SEP-2002; 2002US-0411355P.  
PR 18-SEP-2002; 2002US-0411426P.  
PR 02-OCT-2002; 2002US-0414984P.  
PR 11-OCT-2002; 2002US-0417611P.  
PR 23-OCT-2002; 2002US-0420246P.  
PR 05-NOV-2002; 2002US-0423623P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA (DELZ ) DELTA BIOTECHNOLOGY LTD.  
PA (PRIN-) PRINCIPIA PHARM CORP.  
XX Ballance DJ, Turner AJ, Rosen CA, Haseltine WA;  
XX WPI; 2003-598517/56.  
XX New albumin fusion protein, useful for preparing a composition for  
XX treating diabetes mellitus.  
XX Example 4; SEQ ID NO 279; 24pp; English.  
XX This invention relates to a novel albumin fusion protein having albumin  
XX or biological activity. Human serum albumin is responsible for a  
XX significant proportion of the osmotic pressure of serum and also  
XX functions as a carrier of endogenous and exogenous ligands. The fusion of  
XX albumin to a therapeutic protein may increase shelf-life and stability of  
XX the therapeutic protein. The albumin fusion protein of the invention may  
XX allow production of compositions with antidiabetic activity whilst the  
XX nucleotide sequence which encodes it may be useful for gene therapy. The  
XX albumin fusion protein is useful for preparing a composition for treating  
XX diabetes mellitus. The present sequence is the amino acid sequence of a  
XX novel full-length human albumin therapeutic fusion protein of the  
XX invention. Note: The sequence data for this patent did not form part of  
XX the printed specification, but was obtained in electronic format directly  
XX from WIPO at ftp.wipo.int/pub/publishedpct\_sequences  
XX SQ Sequence 880 AA;

Query Match 83.2%; Score 3151; DB 7; Length 880;

XX	aluminum fusion protein; albumin activity; human serum albumin; serum osmotic pressure; shelf-life; stability; antidiabetic; gene therapy; diabetes mellitus; human.
KW	Chimeric.
OS	Homo sapiens.
XX	WO2003060071-A2.
PN	24-JUL-2003.
PD	23-DEC-2002; 2002WO-US040891.
PF	21-DEC-2001; 2001US-0341811P.
XX	24-JAN-2002; 2002US-0350358P.
PR	28-JAN-2002; 2002US-0351360P.
PR	26-FEB-2002; 2002US-0359370P.
PR	28-FEB-2002; 2002US-0360000P.
PR	27-MAR-2002; 2002US-0367500P.
PR	08-APR-2002; 2002US-0370227P.
PR	10-MAY-2002; 2002US-0378950P.
PR	24-MAY-2002; 2002US-0382617P.
PR	28-MAY-2002; 2002US-0383123P.
PR	05-JUN-2002; 2002US-0385708P.
PR	10-JUL-2002; 2002US-0394625P.
PR	24-JUL-2002; 2002US-0398008P.
PR	13-AUG-2002; 2002US-0402131P.
PR	09-AUG-2002; 2002US-0402708P.
PR	18-SEP-2002; 2002US-0411355P.
PR	18-SEP-2002; 2002US-0411426P.
PR	02-OCT-2002; 2002US-0414984P.
PR	11-OCT-2002; 2002US-0417611P.
PR	23-OCT-2002; 2002US-0420246P.
PR	05-NOV-2002; 2002US-0423623P.
XX	(HUMA-) HUMAN GENOME SCI INC.
PA	(DELZ) DELTA BIOTECHNOLOGY LTD.
PA	(PRIN-) PRINCIPIA PHARM CORP.
XX	Ballance DJ, Turner AJ, Rosen CA, Haseeltine WA;
XX	WPI; 2003-598517/56.
XX	New albumin fusion protein, useful for preparing a composition for treating diabetes mellitus.
PT	Example 4; SEQ ID NO 274; 24pp; English.
XX	This invention relates to a novel albumin fusion protein having albumin or biological activity. Human serum albumin is responsible for a significant proportion of the osmotic pressure of serum and also functions as a carrier of endogenous and exogenous ligands. The fusion of albumin to a therapeutic protein may increase shelf-life and stability of the therapeutic protein. The albumin fusion protein of the invention may allow production of compositions with antidiabetic activity whilst the nucleotide sequence which encodes it may be useful for gene therapy. The albumin fusion protein is useful for preparing a composition for treating diabetes mellitus. The present sequence is the amino acid sequence of a novel full-length human albumin therapeutic fusion protein of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/publishedpct_sequences
XX	Sequence 880 AA;
XX	Query Match 83.2%; Score 3151; DB 7; Length 880;
XX	Best Local Similarity 70.5%; Pred. No. 8.8e-73;
XX	Matches 621; Conservative 17; Mismatches 33; Indels 210; Gaps 11;
Qy	1 DAHKSEVAHRFKDLGEENFKALVLIIFAQYLOQCFFEDHVKLVNEVTEFAKTCVADES
Db	25 DAHKSEVAHRFKDLGEENFKALVLIIFAQYLOQCFFEDHVKLVNEVTEFAKTCVADES

XX albumin fusion protein; albumin activity; human serum albumin;  
KW serum osmotic pressure; shelf-life; stability; antidiabetic;  
KW gene therapy; diabetes mellitus; human.  
XX Chimeric.  
OS Homo sapiens.  
XX WO2003060071-A2.  
XX 24-JUL-2003.  
XX 23-DEC-2002; 2002WO-US040891.  
XX 21-DEC-2001; 2001US-034181P.  
XX 24-JAN-2002; 2002US-0350358P.  
XX 28-JAN-2002; 2002US-0351360P.  
XX 26-FEB-2002; 2002US-0359370P.  
XX 28-FEB-2002; 2002US-036000P.  
XX 27-MAR-2002; 2002US-0367500P.  
XX 08-APR-2002; 2002US-0370227P.  
XX 10-MAY-2002; 2002US-0378950P.  
XX 24-MAY-2002; 2002US-0382617P.  
XX 28-MAY-2002; 2002US-0383123P.  
XX 05-JUN-2002; 2002US-0385708P.  
XX 10-JUL-2002; 2002US-0394625P.  
XX 24-JUL-2002; 2002US-0398008P.  
XX 09-AUG-2002; 2002US-0402131P.  
XX 13-AUG-2002; 2002US-0402708P.  
XX 18-SEP-2002; 2002US-0411355P.  
XX 18-SEP-2002; 2002US-0411426P.  
XX 02-OCT-2002; 2002US-0414984P.  
XX 11-OCT-2002; 2002US-0417611P.  
XX 23-OCT-2002; 2002US-0420246P.  
XX 05-NOV-2002; 2002US-0423623P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA (DELZ ) DELTA BIOTECHNOLOGY LTD.  
PA (PRIN-) PRINCIPIA PHARM CORP.  
XX Ballance DJ, Turner AJ, Rosen CA, Haseltine WA;  
XX WPI; 2003-598517/56.  
XX New albumin fusion protein, useful for preparing a composition for  
XX treating diabetes mellitus.  
XX Example 4; SEQ ID NO 274; 24pp; English.  
XX This invention relates to a novel albumin fusion protein having albumin  
XX or biological activity. Human serum albumin is responsible for a  
XX significant proportion of the osmotic pressure of serum and also  
XX functions as a carrier of endogenous and exogenous ligands. The fusion of  
XX albumin to a therapeutic protein may increase shelf-life and stability of  
XX the therapeutic protein. The albumin fusion protein of the invention may  
XX allow production of compositions with antidiabetic activity whilst the  
XX nucleotide sequence which encodes it may be useful for gene therapy. The  
XX albumin fusion protein is useful for preparing a composition for treating  
XX diabetes mellitus. The present sequence is the amino acid sequence of a  
XX novel full-length human albumin therapeutic fusion protein of the  
XX invention. Note: The sequence data for this patent did not form part of  
XX the printed specification, but was obtained in electronic format directly  
XX from WIPO at ftp.wipo.int/pub/publishedpct\_sequences  
XX Sequence 880 AA;  
XX Query Match 83.2%; Score 3151; DB 7; Length 880;  
XX Best Local Similarity 70.5%; Pred. No. 8.8e-73;  
XX Matches 621; Conservative 17; Mismatches 33; Indels 210; Gaps 11;

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144  
QY 121 DVMCTAFHDNEETFLKKYLVEIARRHPYFYFAPELLFFAKRYKAAFTCCQAADKAACLLP 180  
DB 145 DVMCTAFHDNEETFLKKYLVEIARRHPYFYFAPELLFFAKRYKAAFTCCQAADKAACLLP 204  
QY 181 KLDELREGKASSAKQRLKCSLQKGERAFKAWARLSORFPKAEFAEVSKLVTDLTK 240  
DB 205 KLDELREGKASSAKQRLKCSLQKGERAFKAWARLSORFPKAEFAEVSKLVTDLTK 264  
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKCECKEPLLEKSHCIAEVENDEMPA 300  
DB 265 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKCECKEPLLEKSHCIAEVENDEMPA 324  
QY 301 DLPSLAADFVESKDVCKNYABAKOVFLGMFLFYEARHPDYSVVLLRLAKTYETTTLEK 360  
DB 325 DLPSLAADFVESKDVCKNYABAKOVFLGMFLFYEARHPDYSVVLLRLAKTYETTTLEK 384  
QY 361 CAADPHCEYAKVDEPKPLVEEPQNLIKQNCFLPEQLGEYKFNALLVRYTKVPOVST 420  
DB 385 CAADPHCEYAKVDEPKPLVEEPQNLIKQNCFLPEQLGEYKFNALLVRYTKVPOVST 444  
QY 421 PTLVEVSRLGKVGSKCCCHPEAKRMPCAEEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 445 PTLVEVSRLGKVGSKCCCHPEAKRMPCAEEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 504  
QY 481 LVNRRPCPSALEVDYVYKFNAEFTFFHADICTLSEKERQIKQKQALVELVGHKPKAT 540  
DB 505 LVNRRPCPSALEVDYVYKFNAEFTFFHADICTLSEKERQIKQKQALVELVGHKPKAT 564  
QY 541 KEOLKAVMDDFAAFVEKCKADDDKTCFAESGKLVAAASQAALGLMSPLRLEVP----- 593  
DB 565 KEOLKAVMDDFAAFVEKCKADDDKTCFAESGKLVAAASQAALGLVA---BTPIYPWRDA 621  
QY 594 -----C\$HALPOGLSPQVILVRGLVQE-----PKHFT----- 622  
DB 622 ETGERLCAQCFP-----GTFVQPCRRDSTTCGCPCPRHVYTFQWYLERCR 669  
QY 623 -----VSLRDQAANA----- 632  
DB 670 YCNVLGGEREBEARACHATHNRACRRTGTFPAHAGFCLHASCPPGAGVIAPTGTSQNTQ 729  
QY 633 -----PVTLRAS----- 639  
DB 730 CQPCPPGTFSASSSSSQCPHRNCTALGLALNVPGSSSHDTLCTSGTFPLSTRVPGAE 789  
QY 640 -----FADRTLQWISR-----WG-----QKKLISAPFLFYQRF 668  
DB 790 ECERAVIDFVAFQDISIKELQELQALNAPCGWGTPPRAGRAALQKL-----RRR 840  
QY 669 FEVLLFQEGGLKALNGOGLGATSMNOQALEQLRELISG 709  
DB 841 LTELIGAODGALLVRL-----LQALVARMPG 867

RESULT 26  
ADFI4937  
ID ADFI4937 standard; protein; 880 AA.  
XX  
AC ADFI4937;  
XX  
DT 12-FEB-2004 (first entry)  
XX  
DE Human albumin therapeutic fusion protein SeqID233.  
XX  
KW albumin fusion protein; albumin activity; human serum albumin;  
KW serum osmotic pressure; shelf-life; stability; antidiabetic;  
KW gene therapy; diabetes mellitus; human.  
XX  
OS Chimeric.

OS Homo sapiens.  
PN WO2003060071-A2.  
XX  
PD 24-JUL-2003.  
XX  
PF 23-DEC-2002; 2002WO-US040891.  
XX  
PP 21-DEC-2001; 2001US-0341811P.  
PR 24-JAN-2002; 2002US-0350358P.  
PR 28-JAN-2002; 2002US-0351360P.  
PR 26-FEB-2002; 2002US-0359370P.  
PR 28-FEB-2002; 2002US-0360000P.  
PR 27-MAR-2002; 2002US-0367500P.  
PR 08-APR-2002; 2002US-0370227P.  
PR 10-MAY-2002; 2002US-0378950P.  
PR 24-MAY-2002; 2002US-0382617P.  
PR 28-MAY-2002; 2002US-0383123P.  
PR 05-JUN-2002; 2002US-0385708P.  
PR 10-JUL-2002; 2002US-0394625P.  
PR 24-JUL-2002; 2002US-0398008P.  
PR 09-AUG-2002; 2002US-0402131P.  
PR 13-AUG-2002; 2002US-0402708P.  
PR 18-SEP-2002; 2002US-0411355P.  
PR 18-SEP-2002; 2002US-0411426P.  
PR 02-OCT-2002; 2002US-0414984P.  
PR 11-OCT-2002; 2002US-0417611P.  
PR 23-OCT-2002; 2002US-0420246P.  
PR 05-NOV-2002; 2002US-0423623P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA (DELZ ) DELTA BIOTECHNOLOGY LTD.  
PA (PRIN-) PRINCIPIA PHARM CORP.  
XX  
PI Ballance DJ, Turner AJ, Rosen CA, Haseltine WA;  
XX WPI; 2003-598517/56.  
DR  
XX New albumin fusion protein, useful for preparing a composition for  
XX treating diabetes mellitus.  
PS Example 4; SEQ ID NO 233; 24pp; English.  
XX  
CC This invention relates to a novel albumin fusion protein having albumin  
CC or biological activity. Human serum albumin is responsible for a  
CC significant proportion of the osmotic pressure of serum and also  
CC functions as a carrier of endogenous and exogenous ligands. The fusion of  
CC albumin to a therapeutic protein may increase shelf-life and stability of  
CC the therapeutic protein. The albumin fusion protein of the invention may  
CC allow production of compositions with antidiabetic activity whilst the  
CC nucleotide sequence which encodes it may be useful for gene therapy. The  
CC albumin fusion protein is useful for preparing a composition for treating  
CC diabetes mellitus. The present sequence is the amino acid sequence of a  
CC novel full-length human albumin therapeutic fusion protein of the  
CC invention. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/publishedpct\_sequences  
XX  
SQ Sequence 880 AA;  
Query Match 83.2%; Score 3151; DB 7; Length 880;  
Best Local Similarity 70.5%; Pred. No. 8.8e-73;  
Matches 621; Conservative 17; Mismatches 33; Indels 210; Gaps 11;  
QY 1 DAHSEVAHREFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAG 60  
DB 25 DAHSEVAHREFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAG 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144  
QY 121 DVMCTAFHDNEETFLKKYLVEIARRHPYFYFAPELLFFAKRYKAAFTCCQAADKAACLLP 180

Db 145 DVNCTAFHNEETFLKKYLYEIAARRHPYFYVAPELLFFAKRYKAAFTCCQQAADKAACLLP 204  
Qy 181 KLDELDEGKASSAKORLKASQKFGGERAFKAWARLSORFPKPAFAEFAEVSKLVTDLTK 240  
Db 205 KLDELDEGKASSAKORLKASQKFGGERAFKAWARLSORFPKPAFAEFAEVSKLVTDLTK 264  
Qy 241 VHTTECHGDLLECADRADIAKYICENODSISSKLEKCEKCEKPLEKSHCHIAEVENDEMPA 300  
Db 265 VHTTECHGDLLECADRADIAKYICENODSISSKLEKCEKCEKPLEKSHCHIAEVENDEMPA 324  
Qy 301 DLPSLAADFVESKDVCKNVAEADVFLGMLFYIYARRHPDYSVVLRLRAKTYETTLK 360  
Db 325 DLPSLAADFVESKDVCKNVAEADVFLGMLFYIYARRHPDYSVVLRLRAKTYETTLK 384  
Qy 361 CAAADPHCEYAKVDFPKPIVESPONLIKQNCLEFQOLGEYKFNQALLVRYTKVPQVST 420  
Db 385 CAAADPHCEYAKVDFPKPIVESPONLIKQNCLEFQOLGEYKFNQALLVRYTKVPQVST 444  
Qy 421 PTILVESRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480  
Db 445 PTILVESRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 504  
Qy 481 LVNRRPCFSALEVDYVVPKEFNAETFTFHADICTLSEKEROIKKQATLVKHKPKAT 540  
Db 505 LVNRRPCFSALEVDYVVPKEFNAETFTFHADICTLSEKEROIKKQATLVKHKPKAT 564  
Qy 541 KEOLKAVMDFAAFVEKCKKADKCTCFABEGKLVAAQALGLMSPRLEVP----- 593  
Db 565 KEOLKAVMDFAAFVEKCKKADKCTCFABEGKLVAAQALGLVA---ETPTPWRDA 621  
Qy 594 -----CSHALPQGLSPGVIIVRGLVLOE-----PKHET----- 622  
Db 622 ETGERLVCAOCPP-----GTFVQPCRRDSPTTGCGCPPRHYTFQWNYLERCR 669  
Qy 623 -----VSLRQQAHA----- 632  
Db 670 YCNVLGGEREEARACHATHNACRCHTGFPAHAGFCLHASCPGAGVIAPGTPSQNTQ 729  
Qy 633 -----PVTLRAS----- 639  
Db 730 CQPCPPGTFSSSSSSSQCPHRQCTALGLALNVPGSSSHDTLCTSGTGFPLSTRVFGAE 789  
Qy 640 -----FADRTLQWISR-----WG-----OKKLISAPFLFYPORF 668  
Db 790 ECERAVIDFVAFQDISRIRLQRLQALEAPEGWGTPTPRAGRAALQLKL-----RRR 840  
Qy 669 FEVLLLFQEGGLKIALNGOGLGATSMNQALQLELRISG 709  
Db 841 LTELGLAQDQGALLVRL-----LQALRVARMPG 867

RESULT 27  
ADFL14921  
ID ADF14921 standard; protein; 880 AA.  
XX  
XX  
AC ADF14921;  
XX  
XX  
XX 12-FEB-2004 (first entry)  
XX Human albumin therapeutic fusion protein SeqID217.  
XX albumin fusion protein; albumin activity; human serum albumin;  
XX serum osmotic pressure; shelf-life; stability; antidiabetic;  
XX gene therapy; diabetes mellitus; human.  
XX Chimeric.  
OS Homo sapiens.  
XX  
PN W02003060071-A2.  
XX  
XX 24-JUL-2003.  
PD  
XX

PF 23-DEC-2002; 2002WO-US040891.  
XX  
PR 21-DEC-2001; 2001US-0341811P.  
PR 24-JAN-2002; 2002US-0350358P.  
PR 28-JAN-2002; 2002US-0351360P.  
PR 26-FEB-2002; 2002US-0359370P.  
PR 28-FEB-2002; 2002US-0360000P.  
PR 27-MAR-2002; 2002US-0367500P.  
PR 08-APR-2002; 2002US-0370227P.  
PR 10-MAY-2002; 2002US-0378950P.  
PR 24-MAY-2002; 2002US-0382617P.  
PR 28-MAY-2002; 2002US-0383123P.  
PR 05-JUN-2002; 2002US-0385708P.  
PR 10-JUL-2002; 2002US-0394625P.  
PR 24-JUL-2002; 2002US-0398008P.  
PR 09-AUG-2002; 2002US-0402131P.  
PR 13-AUG-2002; 2002US-0402708P.  
PR 18-SEP-2002; 2002US-0411355P.  
PR 18-SEP-2002; 2002US-0411426P.  
PR 02-OCT-2002; 2002US-0414984P.  
PR 11-OCT-2002; 2002US-0417611P.  
PR 23-OCT-2002; 2002US-0420246P.  
PR 05-NOV-2002; 2002US-0423623P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA (DELZ ) DELTA BIOTECHNOLOGY LTD.  
PA (PRIN-) PRINCIPIA PHARM CORP.  
XX  
PI Ballance DJ, Turner AJ, Rosen CA, Haseltine WA;  
XX WPI; 2003-598517/56.  
DR  
XX New albumin fusion protein, useful for preparing a composition for  
PT treating diabetes mellitus.  
XX  
XX Example 4; SEQ ID NO 217; 24pp; English.  
XX  
XX This invention relates to a novel albumin fusion protein having albumin  
CC or biological activity. Human serum albumin is responsible for a  
CC significant proportion of the osmotic pressure of serum and also  
CC functions as a carrier of endogenous and exogenous ligands. The fusion of  
CC albumin to a therapeutic protein may increase shelf-life and stability of  
CC the therapeutic protein. The albumin fusion protein of the invention may  
CC allow production of compositions with antidiabetic activity whilst the  
CC nucleotide sequence which encodes it may be useful for gene therapy. The  
CC albumin fusion protein is useful for preparing a composition for treating  
CC diabetes mellitus. The present sequence is the amino acid sequence of a  
CC novel full-length human albumin therapeutic fusion protein of the  
CC invention. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/publishedpct\_sequences  
XX  
SQ Sequence 880 AA;  
Query Match 83.2%; Score 3151; DB 7; Length 880;  
Best Local Similarity 70.5%; Pred. No. 8.8e-73;  
Matches 621; Conservative 17; Mismatches 33; Indels 210; Gaps 11;  
Qy 1 DAHSEVAHRFKDLGGENFKALVLIIFAQYLOQCPFDHVKLVNVEYFAKTCVADESAE 60  
Db 25 DAHSEVAHRFKDLGGENFKALVLIIFAQYLOQCPFDHVKLVNVEYFAKTCVADESAE 84  
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGMAADCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
Db 85 NCDKSLHTLFGDKLCTVATLRETYGMAADCAKQEPERNECFLOHKDDNPNLRLVRPEV 144  
Qy 121 DVNCTAFHNEETFLKKYLYEIAARRHPYFYVAPELLFFAKRYKAAFTCCQQAADKAACLLP 180  
Db 145 DVNCTAFHNEETFLKKYLYEIAARRHPYFYVAPELLFFAKRYKAAFTCCQQAADKAACLLP 204  
Qy 181 KLDELDEGKASSAKORLKASQKFGGERAFKAWARLSORFPKPAFAEFAEVSKLVTDLTK 240  
Db 205 KLDELDEGKASSAKORLKASQKFGGERAFKAWARLSORFPKPAFAEFAEVSKLVTDLTK 264

QY 241 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDMPA 300  
DB 265 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDMPA 324  
QY 301 DLPSLAADFVSKDVKCKYAKAKVFLGMFLYEVARRHPDYVSVLLRLAKTYETILLEK 360  
DB 325 DLPSLAADFVSKDVKCKYAKAKVFLGMFLYEVARRHPDYVSVLLRLAKTYETILLEK 384  
QY 361 CAADPHCEYAKVDFEFPKPLVEEQNLKONCELFEOQGEYKFNALLVRYTKVPOVST 420  
DB 385 CAADPHCEYAKVDFEFPKPLVEEQNLKONCELFEOQGEYKFNALLVRYTKVPOVST 444  
QY 421 PTLVEVSNLKGKSGKCKHPEAKRMPCAEDYLSVNLQCLVLEKTPVSDRVTKCTTES 480  
DB 445 PTLVEVSNLKGKSGKCKHPEAKRMPCAEDYLSVNLQCLVLEKTPVSDRVTKCTTES 504  
QY 481 LVNRRPCPSALEVDETVYVPKFNATFTFHADICTLSEKEROIKQOTALVELVGHKPKAT 540  
DB 505 LVNRRPCPSALEVDETVYVPKFNATFTFHADICTLSEKEROIKQOTALVELVGHKPKAT 564  
QY 541 KEQLKAVMDDFAAFVEKCKKADDDKTCFAERGGKLVAAASQAALGLMSPLLEVP----- 593  
DB 565 KEQLKAVMDDFAAFVEKCKKADDDKTCFAERGGKLVAAASQAALGLVA---ETPTYPWRDA 621  
QY 594 -----CCHALPQGLSPGVIIVRGLVLE-----PRHFT----- 622  
DB 622 ETGERLVCAQCP-----GTFTVQPCRRDSTTCGCPPRHYTQFWNLERCR 669  
QY 623 -----VSLRDOAHA----- 632  
DB 670 YCNVLGEBEAREACHATHNRACRCRTGFFAHAGFCLEHASCPPGAGVIAPGTPSQNTQ 729  
QY 633 -----PVTLRAS----- 639  
DB 730 CQPCPPGTFSSSSSEQCPHRNCTALGLALNVGSSSHDTLCTSGFPLSTRVPGA 789  
QY 640 -----FADRTLOWIS-----WG-----QKLLISAPFLFYPORF 668  
DB 790 ECERAVIDFVAFQDISIKRLQRLQALPAEGWGTPPRAGRALQLKL-----RRR 840  
QY 669 FEVLLFQGGKLGALNGQGLGATSMNOQALQLERLRISG 709  
DB 841 LTELGAQDGLLVRL-----LQALRVARMGP 867

RESULT 28  
ADH21282  
ID ADH21282 standard; protein; 880 AA.  
XX ADH21282;  
AC ADH21282;  
XX  
XX 11-MAR-2004 (first entry)  
XX Human albumin/mutant TR6 fusion protein, SEQ ID NO:79.  
XX Fusion protein; human serum albumin; HSA; therapeutic protein;  
KW shelf-life; in vitro biological activity; in vivo biological activity;  
KW metabolic disorder; endocrine disorder; diabetes; type 1; type 2;  
KW diabetes-related condition; hyperglycaemia; neural disorder; neuropathy;  
KW retinopathy; cardiovascular disorder; heart disease; renal disorder;  
KW obesity; glucose level maintenance; weight loss; antidiabetic; cardiant;  
KW anorectic; ophthalmological; gene therapy.  
XX  
XX Synthetic.  
OS Chimeric.  
OS Homo sapiens.  
XX  
XX WO2003059934-A2.  
XX  
XX PD 24-JUL-2003.  
XX  
XX 23-DEC-2002; 2002WO-US040892.

XX 21-DEC-2001; 2001US-0341811P.  
PR 24-JAN-2002; 2002US-0350358P.  
PR 26-FEB-2002; 2002US-0359370P.  
PR 28-FEB-2002; 2002US-0360000P.  
PR 27-MAR-2002; 2002US-0367500P.  
PR 08-APR-2002; 2002US-0370227P.  
PR 10-MAY-2002; 2002US-0378950P.  
PR 24-JUL-2002; 2002US-0398008P.  
PR 09-AUG-2002; 2002US-0402131P.  
PR 13-AUG-2002; 2002US-0402708P.  
PR 18-SEP-2002; 2002US-0411355P.  
PR 02-OCT-2002; 2002US-0414984P.  
PR 11-OCT-2002; 2002US-0417611P.  
PR 23-OCT-2002; 2002US-0420246P.  
PR 05-NOV-2002; 2002US-0423623P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA Rosen CA, Haseltine WA;  
XX  
XX WPI; 2003-598501/56.  
XX  
XX New albumin fusion protein, useful for preparing a composition for  
PT treating diabetes mellitus.  
XX  
XX Disclosure; SEQ ID NO 79; 1086pp; English.  
XX  
XX The invention relates to fusion proteins comprising human serum albumin  
CC (ADH21530) and a therapeutic polypeptide such as a therapeutic protein,  
CC antibody or peptide or their variants or fragments. The therapeutic  
CC protein may be fused to the N-terminus, the C-terminus or both termini of  
CC albumin via a linker. The albumin component of the fusion proteins  
CC prolongs the shelf-life and the in vitro and vivo biological activity of  
CC the proteins compared with those of the corresponding therapeutic  
CC proteins on their own. The invention also relates to nucleic acids  
CC encoding albumin fusion proteins, vectors and host cells comprising an  
CC albumin fusion protein nucleic acid, compositions and kits comprising an  
CC albumin fusion protein, the method of extending the shelf-life of a  
CC therapeutic protein by fusion with albumin, and the treatment of disease  
CC using an albumin fusion protein. The albumin fusion proteins may be used  
CC in the treatment of metabolic/endocrine disorders, diabetes and diabetes-  
CC related conditions. Specifically the albumin fusion proteins may be used  
CC to treat type 1 and type 2 diabetes, hyperglycaemia, neural disorders  
CC (especially neuropathy), retinopathy, cardiovascular disorders  
CC (especially heart disease, renal disorders and obesity). The proteins may  
CC also be used in a method of maintaining a basal glucose level in a  
CC patient and in a method for losing weight. The present sequence is  
CC related to the invention.  
XX  
XX Sequence 880 AA;  
SQ  
Query Match 83.2%; Score 3151; DB 7; Length 880;  
Best Local Similarity 70.5%; Pred. No. 8.8e-73;  
Matches 621; Conservative 17; Mismatches 33; Indels 210; Gaps 11;  
QY 1 DAHSEVAHRFKDLGEENFKALVLIATAFYLIQQCFEDHVKLVNEVTEFAKTCVADESAE 60  
DB 25 DAHSEVAHRFKDLGEENFKALVLIATAFYLIQQCFEDHVKLVNEVTEFAKTCVADESAE 84  
QY 61 NCDKSLHTLFGDKLCTVATLTRETYGEMADCCAKQEPERNECFLOHKDNPMLPLVRPEV 120  
DB 85 NCDKSLHTLFGDKLCTVATLTRETYGEMADCCAKQEPERNECFLOHKDNPMLPLVRPEV 144  
QY 121 DVNCTAFHNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQADKAACLLP 180  
DB 145 DVNCTAFHNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQADKAACLLP 204  
QY 181 KLDELDEGKASSAKQRLKCSAQKFGGERAFKAWAVARLSQRFPAEFAEYKLVTDLT 240  
DB 205 KLDELDEGKASSAKQRLKCSAQKFGGERAFKAWAVARLSQRFPAEFAEYKLVTDLT 264  
QY 241 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDMPA 300



```
Db 265 VHTCECHGDLLECCADDRADLAKYICENQDSISSKLEKCECKPFLKSHCIAEVENDEMPA 324
Qy 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYFYARRHPDYSVVLRLAKTYETTTLEKC 360
Db 325 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYFYARRHPDYSVVLRLAKTYETTTLEKC 384
Qy 361 CAAADPHCEYAKVDFEPKPIVEBPQNLIKONCELFEOLGKYKFNALLVRYTKVPOVST 420
Db 385 CAAADPHCEYAKVDFEPKPIVEBPQNLIKONCELFEOLGKYKFNALLVRYTKVPOVST 444
Qy 421 PTLVEVSRNLGKVGSKCKCHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480
Db 445 PTLVEVSRNLGKVGSKCKCHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 504
Qy 481 LVNRRPCFSALEVDVETVVPKEFNAETFTFHADICTLSEKERQIKQTALVELVHKRPKAT 540
Db 505 LVNRRPCFSALEVDVETVVPKEFNAETFTFHADICTLSEKERQIKQTALVELVHKRPKAT 564
Qy 541 KEOLKAVMDDFAAFVEKCKKADDKETCTFAEEGKLVAAASQAALGLMSPRLEVP----- 593
Db 565 KEOLKAVMDDFAAFVEKCKKADDKETCTFAEEGKLVAAASQAALGLVA---ETPTYPWRDA 621
Qy 594 -----CSHALPQGLSPGQVIVRGLVLOE-----PKHFT----- 622
Db 622 ETGERLVCAQCPP-----GTFVQPCRRDSTPTTCGCPRRHYTQFWNYLERCR 669
Qy 623 -----VSLRQAAVA----- 632
Db 670 YCNVLCGEREEERACHATHNRACRGTGFFAHAGFLEHASCPGAGVIAPGTPSQNTQ 729
Qy 633 -----PVTLRAS----- 639
Db 730 CQPCPPGTFASSSSBQCPHQCTALGLALNVPGSSSHDTLCTSGPPLSTRVPGAE 789
Qy 640 -----FADRTLQWISR-----WG-----QKKLISAPFLFYQRF 668
Db 790 ECERAVIDFVAFQDISIKRLQLLQALEPFGWGPTPRAGRAALQLKL-----RRR 840
Qy 669 FEVILLFOEGGLKIALNGQGLGATSMNQALEQLRELIRISG 709
Db 841 LTELGAQDQALLVRL-----LQALRVARMFG 867

RESULT 29
ADH21273
ID ADH21273 standard; protein; 880 AA.
XX AC ADH21273;
XX DT 11-MAR-2004 (first entry)
XX DE Human albumin/TR6 fragment fusion protein, SEQ ID NO:70.
XX KW Fusion protein; human serum albumin; HSA; therapeutic protein;
XX KW shelf-life; in vitro biological activity; in vivo biological activity;
XX KW metabolic disorder; endocrine disorder; diabetes; type 1; type 2;
XX KW diabetes-related condition; hyperglycaemia; neural disorder; neuropathy;
XX KW retinopathy; cardiovascular disorder; heart disease; renal disorder;
XX KW obesity; glucose level maintenance; weight loss; antidiabetic; cardiant;
XX KW anorectic; ophthalmological; gene therapy.
XX OS Chimeric.
XX OS Homo sapiens.
XX PN WO2003059934-A2.
XX PD 24-JUL-2003.
XX PF 23-DEC-2002; 2002WO-US040892.
XX PR 21-DEC-2001; 2001US-0341811P.
XX PR 24-JAN-2002; 2002US-0350358P.
```

```
PR 26-FEB-2002; 2002US-0359370P.
PR 28-FEB-2002; 2002US-0360000P.
PR 27-MAR-2002; 2002US-0367500P.
PR 08-APR-2002; 2002US-0370227P.
PR 10-MAY-2002; 2002US-0378950P.
PR 24-JUL-2002; 2002US-0398008P.
PR 09-AUG-2002; 2002US-0402131P.
PR 13-AUG-2002; 2002US-0402708P.
PR 18-SEP-2002; 2002US-0411355P.
PR 02-OCT-2002; 2002US-0414984P.
PR 11-OCT-2002; 2002US-0417611P.
PR 23-OCT-2002; 2002US-0420246P.
PR 05-NOV-2002; 2002US-0423623P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX PA
XX PI Rosen CA, Haseltine WA;
XX PT WPI; 2003-598501/56.
XX DR
XX PT New albumin fusion protein, useful for preparing a composition for
XX PT treating diabetes mellitus.
XX PS Disclosure; SEQ ID NO 70; 1086pp; English.
XX CC The invention relates to fusion proteins comprising human serum albumin
XX CC (ADH21530) and a therapeutic polypeptide such as a therapeutic protein,
XX CC antibody or peptide or their variants or fragments. The therapeutic
XX CC protein may be fused to the N-terminus, the C-terminus or both termini of
XX CC albumin via a linker. The albumin component of the fusion proteins
XX CC prolongs the shelf-life and the in vitro and vivo biological activity of
XX CC the proteins compared with those of the corresponding therapeutic
XX CC proteins on their own. The invention also relates to nucleic acids
XX CC encoding albumin fusion proteins, vectors and host cells comprising an
XX CC albumin fusion protein nucleic acid, compositions and kits comprising an
XX CC albumin fusion protein, the method of extending the shelf-life of a
XX CC therapeutic protein by fusion with albumin, and the treatment of disease
XX CC using an albumin fusion protein. The albumin fusion proteins may be used
XX CC in the treatment of metabolic/endocrine disorders, diabetes and diabetes-
XX CC related conditions. Specifically the albumin fusion proteins may be used
XX CC to treat type 1 and type 2 diabetes, hyperglycaemia, neural disorders
XX CC (especially neuropathy), retinopathy, cardiovascular disorders
XX CC (especially heart disease, renal disorders and obesity. The proteins may
XX CC also be used in a method of maintaining a basal glucose level in a
XX CC patient and in a method for losing weight. The present sequence is
XX CC related to the invention.
XX SQ Sequence 880 AA;
Query Match 83.2%; Score 3151; DB 7; Length 880;
Best Local Similarity 70.5%; Pred. No. 8.8e-73;
Matches 621; Conservative 17; Mismatches 33; Indels 210; Gaps 11;
Qy 1 DAKHSEVAHRFKDGLGEENFKALVLIATAQYLQCCPFEDHVKLNVNTEFAKTCVADESAE 60
Db 25 DAKHSEVAHRFKDGLGEENFKALVLIATAQYLQCCPFEDHVKLNVNTEFAKTCVADESAE 84
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPDLRLVRPEV 120
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPDLRLVRPEV 144
Qy 121 DVMCTAFHDNEETFLKLYLYETARRHPYFVAPPELLFFAKRYKAAFTCCQADKAACLIP 180
Db 145 DVMCTAFHDNEETFLKLYLYETARRHPYFVAPPELLFFAKRYKAAFTCCQADKAACLIP 204
Qy 181 KLDELDEGKASSAKQRLKASIQKFGERAFAKAWAVARLSQRFPKAEFVSKLVTDLTK 240
Db 205 KLDELDEGKASSAKQRLKASIQKFGERAFAKAWAVARLSQRFPKAEFVSKLVTDLTK 264
Qy 241 VHTCECHGDLLECCADDRADLAKYICENQDSISSKLEKCECKPFLKSHCIAEVENDEMPA 300
Db 265 VHTCECHGDLLECCADDRADLAKYICENQDSISSKLEKCECKPFLKSHCIAEVENDEMPA 324
```

Qy 301 DLPSLAADFVESKDVCKNYABAKVFLGMFLYFYEARHDPYVSVLLRLAKTYTTLEKC 360  
Db 325 DLPSLAADFVESKDVCKNYABAKVFLGMFLYFYEARHDPYVSVLLRLAKTYTTLEKC 384  
Qy 361 CAADPHCYAKVDEKPLVEEPPONLTKONCELFQELGEYKFNALLVRYTKVPOVST 420  
Db 385 CAADPHCYAKVDEKPLVEEPPONLTKONCELFQELGEYKFNALLVRYTKVPOVST 444  
Qy 421 PTLVEVRNLTGKVGSKCKCHPEAKRMPCAEYDLSVNLNQLCVLHEKTPVSDRVTKCCTES 480  
Db 445 PTLVEVRNLTGKVGSKCKCHPEAKRMPCAEYDLSVNLNQLCVLHEKTPVSDRVTKCCTES 504  
Qy 481 LVNRRPCFSALEVEDTYVPKEFNAETFTTHADICTLSEKEROIKQOTALVELVGHKPKAT 540  
Db 505 LVNRRPCFSALEVEDTYVPKEFNAETFTTHADICTLSEKEROIKQOTALVELVGHKPKAT 564  
Qy 541 KEQLKAVNDDFAAFVEKCKCKADDDKTCFAEKGKLVAAQAALGLMSPRLVPP----- 593  
Db 565 KEQLKAVNDDFAAFVEKCKCKADDDKTCFAEKGKLVAAQAALGLVA---ETPTYPWDA 621  
Qy 594 -----CSHALPQGLSPGVIIVRGLVLQE-----PKHFT----- 622  
Db 622 ETGERLCAQCP-----GTFVQPCRRSDPTTCGPCPRHYTQFWNLERCR 669  
Qy 623 -----VSLRDOAHV----- 632  
Db 670 YCNVLCGSEARACHATHNRCRCRTGFFAHAGFCLEHASCPCPGAGVAPGTPSQNTQ 729  
Qy 633 -----PVTLRAS----- 639  
Db 730 CQCPGPTFSASSSSSEQCQPHRNTALGLALNVGSSSHDTLCTSGTGFPLSTRVGA 789  
Qy 640 -----FADTLQWISR-----WG-----QKKLISAPFLFYPORF 668  
Db 790 ECERAVIDFVAFQDISIKRLQRLQALRAPEGWGTTPAGRAALQLKL-----RRR 840  
Qy 669 FEVLLFQEGGLKALNGOGLGATSMNQOALEQRLRISG 709  
Db 841 LTELGAQDGLLVRL-----LQALRVARMPG 867

RESULT 30  
ADH21293  
ID ADH21293 standard; protein; 880 AA.  
XX ADH21293;  
AC ADH21293;  
XX 11-MAR-2004 (first entry)  
DT Human albumin/TR6 fragment fusion protein, SEQ ID NO:90.  
DE  
XX Fusion protein; human serum albumin; HSA; therapeutic protein;  
KW shelf-life; in vitro biological activity; in vivo biological activity;  
KW metabolic disorder; endocrine disorder; diabetes; type 1; type 2;  
KW diabetes-related condition; hyperglycaemia; neural disorder; neuropathy;  
KW retinopathy; cardiovascular disorder; heart disease; renal disorder;  
KW obesity; glucose level maintenance; weight loss; antidiabetic; cardiant;  
KW anorectic; ophthalmological; gene therapy.  
XX  
OS Chimeric.  
OS Homo sapiens.  
PN WO2003059934-A2.  
XX 24-JUL-2003.  
PD  
XX 23-DEC-2002; 2002WO-US040892.  
XX  
XX 21-DEC-2001; 2001US-0341811P.  
PR 24-JAN-2002; 2002US-0350358P.  
PR 26-FEB-2002; 2002US-0359370P.  
PR 28-FEB-2002; 2002US-0360000P.  
PR 27-MAR-2002; 2002US-0367500P.

PR 08-APR-2002; 2002US-0370227P.  
PR 10-MAY-2002; 2002US-0378950P.  
PR 24-JUL-2002; 2002US-0398008P.  
PR 09-AUG-2002; 2002US-0402131P.  
PR 13-AUG-2002; 2002US-0402708P.  
PR 18-SEP-2002; 2002US-0411355P.  
PR 02-OCT-2002; 2002US-0414984P.  
PR 23-OCT-2002; 2002US-0417611P.  
PR 05-NOV-2002; 2002US-0420246P.  
PR (HUMA-) HUMAN GENOME SCI INC.  
PA  
XX Rosen CA, Haseltine WA;  
XX WPI; 2003-598501/56.  
XX  
XX New albumin fusion protein, useful for preparing a composition for  
PT treating diabetes mellitus.  
XX  
XX Disclosure; SEQ ID NO 90; 1086pp; English.  
XX  
XX The invention relates to fusion proteins comprising human serum albumin  
CC (ADH21530) and a therapeutic polypeptide such as a therapeutic protein,  
CC antibody or peptide or their variants or fragments. The therapeutic  
CC protein may be fused to the N-terminus, the C-terminus or both termini of  
CC albumin via a linker. The albumin component of the fusion proteins  
CC prolongs the shelf-life and the in vitro and vivo biological activity of  
CC the proteins compared with those of the corresponding therapeutic  
CC proteins on their own. The invention also relates to nucleic acids  
CC encoding albumin fusion proteins, vectors and host cells comprising an  
CC albumin fusion protein nucleic acid, compositions and kits comprising an  
CC albumin fusion protein, the method of extending the shelf-life of a  
CC therapeutic protein by fusion with albumin, and the treatment of disease  
CC using an albumin fusion protein. The albumin fusion proteins may be used  
CC in the treatment of metabolic/endocrine disorders, diabetes and diabetes-  
CC related conditions. Specifically the albumin fusion proteins may be used  
CC to treat type 1 and type 2 diabetes, hyperglycaemia, neural disorders  
CC (especially neuropathy), retinopathy, cardiovascular disorders  
CC (especially heart disease, renal disorders and obesity). The proteins may  
CC also be used in a method of maintaining a basal glucose level in a  
CC patient and in a method for losing weight. The present sequence is  
CC related to the invention.  
XX  
XX Sequence 880 AA;  
SQ  
Query Match 83.2%; Score 3151; DB 7; Length 880;  
Best Local Similarity 70.5%; Pred. No. 8.8e-73;  
Matches 621; Conservative 17; Mismatches 33; Indels 210; Gaps 11;  
Qy 1 DAHSEVAHRFKDILGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
Db 25 DAHSEVAHRFKDILGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 84  
Qy 61 NCDKSLHTLFGDKLCTVATLRETTGEMADCCAKOEPERNECFLOHKDNPMLRLVREV 120  
Db 85 NCDKSLHTLFGDKLCTVATLRETTGEMADCCAKOEPERNECFLOHKDNPMLRLVREV 144  
Qy 121 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQADRAACLLP 180  
Db 145 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQADRAACLLP 204  
Qy 181 KLDELDEGKASSAKORLKCAASLOKFGERAFAKAWAVARLSQRFPAEFAVSKLVTDLTK 240  
Db 205 KLDELDEGKASSAKORLKCAASLOKFGERAFAKAWAVARLSQRFPAEFAVSKLVTDLTK 264  
Qy 241 VHTCCCHGDLLECAADDRADLAKYICENQDSISSKLECCCEKPALEKSHCIAEVENDEMPA 300  
Db 265 VHTCCCHGDLLECAADDRADLAKYICENQDSISSKLECCCEKPALEKSHCIAEVENDEMPA 324  
Qy 301 DLPSLAADFVESKDVCKNYABAKVFLGMFLYFYEARHDPYVSVLLRLAKTYTTLEKC 360  
Db 325 DLPSLAADFVESKDVCKNYABAKVFLGMFLYFYEARHDPYVSVLLRLAKTYTTLEKC 384

```

Qy 361 CAADPHECYAKVDFEKPILVERPQNLIKONCELFEOLGEYKFEONALLVRYTKKVPQVST 420
Db 385 CAADPHECYAKVDFEKPILVERPQNLIKONCELFEOLGEYKFEONALLVRYTKKVPQVST 444
Qy 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVNLQNLVLHEKTPVSDRVTKCTES 480
Db 445 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVNLQNLVLHEKTPVSDRVTKCTES 504
Qy 481 LVNRPCFSALEVDETVVPKEFNAETFTPHADICTLSEKERQIKKOTALVELVKHKPKAT 540
Db 505 LVNRPCFSALEVDETVVPKEFNAETFTPHADICTLSEKERQIKKOTALVELVKHKPKAT 564
Qy 541 KEOLKAVMDDFAAFEVKCKCCKADDDKCTCFABEGKKLVAAQAALGLMSPRLEVP----- 593
Db 565 KEOLKAVMDDFAAFEVKCKCCKADDDKCTCFABEGKKLVAAQAALGLVA---ETPTYPMRDA 621
Qy 594 -----CShALPQGLSPGQVITVRGLVLQE-----PKHFT----- 622
Db 622 ETGERLVCAQCPP-----GTFVORPCERDSPTTCGPPRHHYTOFWNYLERCR 669
Qy 623 -----VSLRDOAHA----- 632
Db 670 YCNVLCGEREEERACHATHNRACRRTGPFPAHAGFCLHASCPCPGAGVIAPGTPSQNTQ 729
Qy 633 -----PVTLRAS----- 639
Db 730 CQPCPPGTFSSSSSSSQCPHRNCTALGLALNVGSSSHDTLCTCTGTPPLSTRVPGAE 789
Qy 640 -----FADRTLQWISR-----WG-----QKKLISAPELFYPORF 668
Db 790 ECERAVIDFVAFODISIKRLQRLLOALEAPEGNGPTPRAGRAALQLKL-----RRR 840
Qy 669 FEVILLFQEGGLKALNGGLGATSMNQALQLEOLRELIRSG 709
Db 841 LTELGAQDQGALLVRL-----LQALRVARMFG 867

```

Search completed: October 17, 2005, 08:29:01  
 Job time : 180 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2005, 08:25:57 ; Search time 43 Seconds  
(without alignments)  
1246.466 Million cell updates/sec

Title: US-10-933-523-18  
Perfect score: 3785  
Sequence: 1 DAHKEVAHRFKDLGERNFK.....leglreirsgvqlycvhs 718

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.1

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Issued Patents\_AA\*

- 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/iaa/PTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3140.6	83.0	783	1	US-08-256-938-2
2	3114.9	82.3	978	4	US-08-897-956A-3
3	3103	82.0	585	1	US-08-153-799-14
4	3103	82.0	585	2	US-08-702-572-2
5	3103	82.0	585	3	US-08-769-746-2
6	3103	82.0	609	4	US-09-976-594-977
7	3103	82.0	609	4	US-09-919-039-370
8	3103	82.0	610	2	US-08-797-689-2
9	3103	82.0	610	4	US-09-984-186-2
10	3103	82.0	622	4	US-09-949-016-11170
11	3103	82.0	787	1	US-08-256-938-4
12	3103	82.0	787	2	US-08-797-689-16
13	3103	82.0	787	4	US-09-984-186-16
14	3099	81.9	609	1	US-08-222-619-3
15	3099	81.9	609	1	US-08-433-037-4
16	3099	81.9	609	4	US-08-897-956A-2
17	3099	81.9	509	5	PCT-US95-04075-3
18	3093	81.7	585	1	US-08-448-196A-3
19	3093	81.7	585	2	US-08-984-176-1
20	2458.9	65.0	583	1	US-08-448-196A-5
21	2450.9	64.8	583	1	US-08-448-196A-4
22	2432.9	64.3	583	1	US-08-448-196A-6
23	2426	64.1	584	1	US-08-448-196A-7
24	2389.8	63.1	582	1	US-08-134-638-1
25	1252.7	33.1	609	1	US-08-222-619-5
26	1252.7	33.1	609	4	US-09-976-594-456
27	1252.7	33.1	609	5	PCT-US95-04075-4

28	1252.7	33.1	612	4	US-09-949-016-11201	Sequence 11201, A
29	1211.3	32.0	609	4	US-09-186-949A-2	Sequence 2, Appli
30	1210.9	32.0	590	2	US-08-377-309-2	Sequence 2, Appli
31	1210.9	32.0	590	3	US-09-186-723-2	Sequence 2, Appli
32	1210.9	32.0	590	3	US-08-505-012-5	Sequence 5, Appli
33	1210.9	32.0	590	4	US-09-186-949A-3	Sequence 3, Appli
34	1210.9	32.0	590	4	US-08-758-757-2	Sequence 2, Appli
35	1210.9	32.0	590	4	US-09-187-378-2	Sequence 2, Appli
36	1210.9	32.0	590	4	US-10-115-701A-2	Sequence 2, Appli
37	1210.9	32.0	590	4	US-09-940-308A-2	Sequence 2, Appli
38	1210.9	32.0	590	4	US-09-940-308A-2	Sequence 2, Appli
39	1210.9	32.0	590	5	PCT-US96-00996-5	Sequence 5, Appli
40	1168.1	30.9	579	1	US-08-448-196A-8	Sequence 8, Appli
41	1059	28.0	599	1	US-08-222-619-2	Sequence 2, Appli
42	1059	28.0	599	3	US-08-221-767-24	Sequence 24, Appli
43	1059	28.0	599	5	PCT-US95-04075-2	Sequence 2, Appli
44	926	24.5	393	2	US-08-377-309-7	Sequence 7, Appli
45	926	24.5	393	3	US-09-186-723-7	Sequence 7, Appli
46	926	24.5	393	3	US-08-505-012-10	Sequence 10, Appli
47	926	24.5	393	4	US-09-186-949A-8	Sequence 8, Appli
48	926	24.5	393	4	US-08-758-757-7	Sequence 7, Appli
49	926	24.5	393	4	US-09-187-978-7	Sequence 7, Appli
50	926	24.5	393	4	US-10-115-701A-7	Sequence 7, Appli
51	926	24.5	393	4	US-09-940-308A-7	Sequence 7, Appli
52	926	24.5	393	4	US-09-940-308A-7	Sequence 7, Appli
53	926	24.5	393	5	PCT-US96-00996-10	Sequence 10, Appli
54	777	20.5	324	3	US-08-505-012-12	Sequence 12, Appli
55	777	20.5	324	5	PCT-US96-00996-12	Sequence 12, Appli
56	777	20.5	325	2	US-08-377-309-8	Sequence 8, Appli
57	777	20.5	325	3	US-09-186-723-8	Sequence 8, Appli
58	777	20.5	325	3	US-08-505-012-11	Sequence 11, Appli
59	777	20.5	325	4	US-09-186-949A-9	Sequence 9, Appli
60	777	20.5	325	4	US-08-758-757-8	Sequence 8, Appli
61	777	20.5	325	4	US-09-187-978-8	Sequence 8, Appli
62	777	20.5	325	4	US-10-115-701A-8	Sequence 8, Appli
63	777	20.5	325	4	US-09-940-308A-8	Sequence 8, Appli
64	777	20.5	325	5	US-09-940-308A-8	Sequence 8, Appli
65	765.3	20.2	590	1	PCT-US96-00996-11	Sequence 11, Appli
66	709.5	18.7	336	3	US-08-448-196A-9	Sequence 9, Appli
67	708.3	18.7	275	4	US-09-131-648-1	Sequence 1, Appli
68	704.9	18.6	286	4	US-09-557-170A-25	Sequence 25, Appli
69	688.9	18.2	389	2	US-08-377-309-6	Sequence 27, Appli
70	688.9	18.2	389	3	US-09-186-723-6	Sequence 6, Appli
71	688.9	18.2	389	3	US-08-505-012-9	Sequence 9, Appli
72	688.9	18.2	389	4	US-09-186-949A-7	Sequence 7, Appli
73	688.9	18.2	389	4	US-08-758-757-6	Sequence 6, Appli
74	688.9	18.2	389	4	US-09-187-378-6	Sequence 6, Appli
75	688.9	18.2	389	4	US-10-115-701A-6	Sequence 6, Appli
76	688.9	18.2	389	4	US-09-940-308A-6	Sequence 6, Appli
77	688.9	18.2	389	4	US-09-940-308A-6	Sequence 6, Appli
78	688.9	18.2	389	4	US-09-940-308A-6	Sequence 6, Appli
79	688.9	18.2	389	5	PCT-US96-00996-9	Sequence 9, Appli
80	676	17.9	133	4	US-09-557-170A-2	Sequence 2, Appli
81	523	13.8	201	2	US-08-377-309-5	Sequence 5, Appli
82	523	13.8	201	3	US-09-186-723-5	Sequence 5, Appli
83	523	13.8	201	3	US-08-505-012-8	Sequence 8, Appli
84	523	13.8	201	4	US-09-186-949A-6	Sequence 6, Appli
85	523	13.8	201	4	US-08-758-757-5	Sequence 5, Appli
86	523	13.8	201	4	US-09-187-378-5	Sequence 5, Appli
87	523	13.8	201	4	US-10-115-701A-5	Sequence 5, Appli
88	523	13.8	201	4	US-09-940-308A-5	Sequence 5, Appli
89	523	13.8	201	4	US-09-940-308A-5	Sequence 5, Appli
90	523	13.8	201	5	PCT-US96-00996-8	Sequence 8, Appli
91	521	13.8	117	1	US-08-448-196A-2	Sequence 2, Appli
92	516	13.6	109	1	US-08-448-196A-1	Sequence 1, Appli
93	421	11.1	474	4	US-09-949-016-11630	Sequence 11630, A
94	417.5	11.0	458	6	5177002-2	Patent No. 5177002
95	417.5	11.0	474	1	US-08-222-619-5	Sequence 5, Appli
96	417.5	11.0	474	5	PCT-US95-04075-5	Sequence 5, Appli
97	417.5	11.0	458	4	US-08-618-4858-1	Sequence 1, Appli
98	416	11.0	458	6	5177002-1	Patent No. 5177002
99	416	11.0	458	6	5177002-1	Patent No. 5177002
100	416	11.0	458	6	5177002-1	Patent No. 5177002



181 KLDLDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRFPAKAEVSKLVTLTK 240  
181 KLDLDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRFPAKAEVSKLVTLTK 451  
241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKCECKPILKSHCIAEVENDEMPA 300  
241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKCECKPILKSHCIAEVENDEMPA 511  
301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYIARRHPDYISVLLRLAKTYETTLK 360  
512 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYIARRHPDYISVLLRLAKTYETTLK 571  
361 CAAADPHECYAKVDFEFPKPLVEEPQNLIKQNCSELFQKGEYKFNALLVRYTKVPQVST 420  
572 CAAADPHECYAKVDFEFPKPLVEEPQNLIKQNCSELFQKGEYKFNALLVRYTKVPQVST 631  
421 PTLVEVSRLGKVGSKCKKHPKAEKMPCAEDYLSVNLQCLVLEKTPVSDRVTKCTTES 480  
632 PTLVEVSRLGKVGSKCKKHPKAEKMPCAEDYLSVNLQCLVLEKTPVSDRVTKCTTES 691  
481 LVNRRPCFSALEVDYVYKPEFNAETFTFHADICTLSEKEROIKQTALVELVHKPKAT 540  
692 LVNRRPCFSALEVDYVYKPEFNAETFTFHADICTLSEKEROIKQTALVELVHKPKAT 751  
541 KEQLKAVMDDFAAFVEKCKADDKETCFPAEGKKLVAASQAALGMSRLEVPCHALPQ 600  
752 KEQLKAVMDDFAAFVEKCKADDKETCFPAEGKKLVAASQAALGMSRLEVPCHALPQ 795  
601 GLSPGOVIIVGLVLOEPKHTVSLRDOAAHAP-----VTIRASPADRTLOWISRW 651  
796 -----GGGSPVKPK-----VSL-----NPPWNRIFKGENVTLCN----- 826  
652 GQKQLISAPPLFYPPRPEV--LLLFBQGGKLALNGOGLGATSMNQALPOLRELRTSG 709  
827 -----GNNFEVSTKWFHNGSLSEETN-----SSLN-----IVNAKPED 861  
710 SVQLYCVH 717  
862 SGEYKQ 869

RESULT 3  
US-08-153-799-14  
; Sequence 14, Application US/08153799  
; Patent No. 5766883  
; GENERAL INFORMATION:  
; APPLICANT: Ballance, David J  
; APPLICANT: Goodey, Andrew R  
; TITLE OF INVENTION: Polypeptides  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: R Hain Swope, BOC Health Care Inc  
; STREET: 100 Mountain Avenue  
; CITY: Murray Hill  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07974  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/153,799  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/847975  
; FILING DATE: 06-MAR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 8909916.2  
; FILING DATE: 29-APR-1989  
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/GB90/00650  
FILING DATE: 26-APR-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/775952  
FILING DATE: 29-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Swope, R Hain  
REGISTRATION NUMBER: 24864  
REFERENCE/DOCKET NUMBER: 92H832  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 665 2400  
TELEFAX: (908) 771 6159  
TELEX: 219484  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 585 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: Region  
LOCATION: 369..419  
OTHER INFORMATION: /note= "Alternative C-termini of  
OTHER INFORMATION: HSA(1-n)"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 1..585  
OTHER INFORMATION: /note= "Amino acid sequence of  
OTHER INFORMATION: natural HSA"  
US-08-153-799-14  
Query Match 82.0%; Score 3103; DB 1; Length 585;  
Best Local Similarity 100.0%; Pred. No. 3,9e-84;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 DAHSEVAHRFKDILGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
Db 1 DAHSEVAHRFKDILGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOPEPERNECFLOHKDNDNPLRLVRPEV 120  
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOPEPERNECFLOHKDNDNPLRLVRPEV 120  
Qy 121 DVMCTAFHDNEETFLKKYLVEIARRHPYVFAPELLFFAKRYKAAAFTECCQAADKAAACLLP 180  
Db 121 DVMCTAFHDNEETFLKKYLVEIARRHPYVFAPELLFFAKRYKAAAFTECCQAADKAAACLLP 180  
Qy 181 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRFPAKAEVSKLVTLTK 240  
Db 181 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRFPAKAEVSKLVTLTK 240  
Qy 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKCECKPILKSHCIAEVENDEMPA 300  
Db 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKCECKPILKSHCIAEVENDEMPA 300  
Qy 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYIARRHPDYISVLLRLAKTYETTLK 360  
Db 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYIARRHPDYISVLLRLAKTYETTLK 360  
Qy 361 CAAADPHECYAKVDFEFPKPLVEEPQNLIKQNCSELFQKGEYKFNALLVRYTKVPQVST 420  
Db 361 CAAADPHECYAKVDFEFPKPLVEEPQNLIKQNCSELFQKGEYKFNALLVRYTKVPQVST 420  
Qy 421 PTLVEVSRLGKVGSKCKKHPKAEKMPCAEDYLSVNLQCLVLEKTPVSDRVTKCTTES 480  
Db 421 PTLVEVSRLGKVGSKCKKHPKAEKMPCAEDYLSVNLQCLVLEKTPVSDRVTKCTTES 480  
Qy 481 LVNRRPCFSALEVDYVYKPEFNAETFTFHADICTLSEKEROIKQTALVELVHKPKAT 540  
Db 481 LVNRRPCFSALEVDYVYKPEFNAETFTFHADICTLSEKEROIKQTALVELVHKPKAT 540





```
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Qy 121 DVMTAFHDNEETFLKKYLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQAAADKAACLLP 180
Db 121 DVMTAFHDNEETFLKKYLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQAAADKAACLLP 180
Qy 181 KLDELDEGKASSAKQRLKCSAQKFGERAFKAWAVARLSQRPFAEFAEVSCLVTDLTk 240
Db 181 KLDELDEGKASSAKQRLKCSAQKFGERAFKAWAVARLSQRPFAEFAEVSCLVTDLTk 240
Qy 241 VHTTECHGDLLECCADRADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVENDEMPA 300
Db 241 VHTTECHGDLLECCADRADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVENDEMPA 300
Qy 301 DLPSLAADFVESKDVCKNVAEAKDVLGFMFLYFYARRHPDYSVVLLRLAKTYETTLLEK 360
Db 301 DLPSLAADFVESKDVCKNVAEAKDVLGFMFLYFYARRHPDYSVVLLRLAKTYETTLLEK 360
Qy 361 CAAADPHECYAKVDFEKFPLVEBPQNLIKQNCLEFQOLGEYKFNQALLVRYTKKVPQVST 420
Db 361 CAAADPHECYAKVDFEKFPLVEBPQNLIKQNCLEFQOLGEYKFNQALLVRYTKKVPQVST 420
Qy 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
Qy 481 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVKHKPKAT 540
Db 481 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVKHKPKAT 540
Qy 541 KEQLKAVMDDFAAVFEKCKCKADDDKTCFAEEGKKLVAASQAALGL 585
Db 541 KEQLKAVMDDFAAVFEKCKCKADDDKTCFAEEGKKLVAASQAALGL 585

RESULT 6
US-09-976-594-977
; Sequence 977, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 977
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 088957CD1
US-09-976-594-977

Query Match 82.0%; Score 3103; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 4.2e-84;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHRFKDLGEENFKALVLIAPQYLQCCPFEDHVKLVNVEVTEFAKTCVADESAAE 60
Db 25 DAHSEVAHRFKDLGEENFKALVLIAPQYLQCCPFEDHVKLVNVEVTEFAKTCVADESAAE 84
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144
Qy 121 DVMTAFHDNEETFLKKYLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQAAADKAACLLP 180
Db 121 DVMTAFHDNEETFLKKYLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQAAADKAACLLP 180
Qy 181 KLDELDEGKASSAKQRLKCSAQKFGERAFKAWAVARLSQRPFAEFAEVSCLVTDLTk 240
Db 181 KLDELDEGKASSAKQRLKCSAQKFGERAFKAWAVARLSQRPFAEFAEVSCLVTDLTk 240
```

```
Db 145 DVMTAFHDNEETFLKKYLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQAAADKAACLLP 204
Qy 181 KLDELDEGKASSAKQRLKCSAQKFGERAFKAWAVARLSQRPFAEFAEVSCLVTDLTk 240
Db 205 KLDELDEGKASSAKQRLKCSAQKFGERAFKAWAVARLSQRPFAEFAEVSCLVTDLTk 264
Qy 241 VHTTECHGDLLECCADRADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVENDEMPA 300
Db 265 VHTTECHGDLLECCADRADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVENDEMPA 324
Qy 301 DLPSLAADFVESKDVCKNVAEAKDVLGFMFLYFYARRHPDYSVVLLRLAKTYETTLLEK 360
Db 325 DLPSLAADFVESKDVCKNVAEAKDVLGFMFLYFYARRHPDYSVVLLRLAKTYETTLLEK 384
Qy 361 CAAADPHECYAKVDFEKFPLVEBPQNLIKQNCLEFQOLGEYKFNQALLVRYTKKVPQVST 420
Db 385 CAAADPHECYAKVDFEKFPLVEBPQNLIKQNCLEFQOLGEYKFNQALLVRYTKKVPQVST 444
Qy 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 445 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 504
Qy 481 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVKHKPKAT 540
Db 505 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVKHKPKAT 564
Qy 541 KEQLKAVMDDFAAVFEKCKCKADDDKTCFAEEGKKLVAASQAALGL 585
Db 565 KEQLKAVMDDFAAVFEKCKCKADDDKTCFAEEGKKLVAASQAALGL 609

RESULT 7
US-09-919-039-370
; Sequence 370, Application US/09919039
; Patent No. 6727066
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT APPLICATION NUMBER: US/09/919,039
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/222,113
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 370
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6727066 088957CD1
US-09-919-039-370

Query Match 82.0%; Score 3103; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 4.2e-84;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHRFKDLGEENFKALVLIAPQYLQCCPFEDHVKLVNVEVTEFAKTCVADESAAE 60
Db 25 DAHSEVAHRFKDLGEENFKALVLIAPQYLQCCPFEDHVKLVNVEVTEFAKTCVADESAAE 84
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144
Qy 121 DVMTAFHDNEETFLKKYLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQAAADKAACLLP 180
Db 145 DVMTAFHDNEETFLKKYLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQAAADKAACLLP 204
Qy 181 KLDELDEGKASSAKQRLKCSAQKFGERAFKAWAVARLSQRPFAEFAEVSCLVTDLTk 240
Db 205 KLDELDEGKASSAKQRLKCSAQKFGERAFKAWAVARLSQRPFAEFAEVSCLVTDLTk 264
```

```
QY 241 VHTCCGDLLECCADRADLAKYICENODSISSKLKECCCKPILLESKSHCIAEVENDEMPA 300
DB 265 VHTCCGDLLECCADRADLAKYICENODSISSKLKECCCKPILLESKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLEYARRHPDYSVVLRLRLAKTYETTTLEKC 360
DB 325 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLEYARRHPDYSVVLRLRLAKTYETTTLEKC 384
QY 361 CAAADPHECYAKVDFEFKPLVEBPQNLIKQNCLEFQOLGEYKFPQNALVRYTKVPQVST 420
DB 385 CAAADPHECYAKVDFEFKPLVEBPQNLIKQNCLEFQOLGEYKFPQNALVRYTKVPQVST 444
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 445 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTLALVELVKHKPKAT 540
DB 505 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTLALVELVKHKPKAT 564
QY 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 585
DB 565 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 609

RESULT 8
US-08-797-689-2
; Sequence 2, Application US/08797689
; Patent No. 5876969
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; APPLICANT: Fournier, Alain
; APPLICANT: Guittou, Jean-Dominique
; APPLICANT: Jung, Gerard
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (Patentin)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; PRIOR APPLICATION DATA: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
```

```
; LENGTH: 610 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-797-689-2

Query Match 82.0%; Score 3103; DB 2; Length 610;
Best Local Similarity 100.0%; Pred. No. 4.2e-84;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSVAHRFKDLGEENFKALVLIATAFYLOQCPPEHVKLVNVEYTERAKTCVADESAE 60
DB 25 DAHKSVAHRFKDLGEENFKALVLIATAFYLOQCPPEHVKLVNVEYTERAKTCVADESAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPNERCEFLQHKDDNPRLVRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPNERCEFLQHKDDNPRLVRPEV 144
QY 121 DVNCTAFHDNEETFLKKLYEYARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAAACLLP 180
DB 145 DVNCTAFHDNEETFLKKLYEYARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAAACLLP 204
QY 181 KLDELDEGKASSAKORLKASLQKFGERAFAKAWAVARLSORPPKAEPAEVSKLVDLT 240
DB 205 KLDELDEGKASSAKORLKASLQKFGERAFAKAWAVARLSORPPKAEPAEVSKLVDLT 264
QY 241 VHTCCGDLLECCADRADLAKYICENODSISSKLKECCCKPILLESKSHCIAEVENDEMPA 300
DB 265 VHTCCGDLLECCADRADLAKYICENODSISSKLKECCCKPILLESKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLEYARRHPDYSVVLRLRLAKTYETTTLEKC 360
DB 325 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLEYARRHPDYSVVLRLRLAKTYETTTLEKC 384
QY 361 CAAADPHECYAKVDFEFKPLVEBPQNLIKQNCLEFQOLGEYKFPQNALVRYTKVPQVST 420
DB 385 CAAADPHECYAKVDFEFKPLVEBPQNLIKQNCLEFQOLGEYKFPQNALVRYTKVPQVST 444
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 445 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTLALVELVKHKPKAT 540
DB 505 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTLALVELVKHKPKAT 564
QY 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 585
DB 565 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 609

RESULT 9
US-09-984-186-2
; Sequence 2, Application US/09984186
; Patent No. 6686179
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; APPLICANT: Fournier, Alain
; APPLICANT: Guittou, Jean-Dominique
; APPLICANT: Jung, Gerard
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
```

MEDIUM TYPE: Floppy disk  
COMPUTER: Macintosh  
OPERATING SYSTEM: System 7.1  
SOFTWARE: Word 5.1 (PatentIn)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/984,186  
FILING DATE: 29-Oct-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/797,689  
FILING DATE: 31-JAN-1997  
APPLICATION NUMBER: US 08/256,927  
FILING DATE: 28-JUL-1994  
APPLICATION NUMBER: FR 92/01064  
FILING DATE: 31-JAN-1992  
APPLICATION NUMBER: PCT/FR93/00085  
FILING DATE: 28-JAN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith P.D., Julie K.  
REGISTRATION NUMBER: P-38,619  
REFERENCE/DOCKET NUMBER: ST92006-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (610) 454-3839  
TELEFAX: (610) 454-3808  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 610 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-984-186-2

Query Match 82.0%; Score 3103; DB 4; Length 610;  
Best Local Similarity 100.0%; Pred. No. 4.2e-84;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHFKDGLGEENFKALVLI AFAQYLQQCPFFEDHVKLVNEVTEFAKTCVADSEAE 60  
Db 25 DAHSEVAHFKDGLGEENFKALVLI AFAQYLQQCPFFEDHVKLVNEVTEFAKTCVADSEAE 84

Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144

Qy 121 DVMCTAFHDNEETFLKKLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQADKAACLLP 180  
Db 145 DVMCTAFHDNEETFLKKLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQADKAACLLP 204

Qy 181 KLDELDEGKASSAKORLKASLQKFGERAFAKAWAVARLSORFPKAFPAEVS KLVTDLTK 240  
Db 205 KLDELDEGKASSAKORLKASLQKFGERAFAKAWAVARLSORFPKAFPAEVS KLVTDLTK 264

Qy 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 300  
Db 265 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 324

Qy 301 DLPSLAADFVESKDVCKNYAEAKOVFLGMFLYFYARRHPDYSVLLRLAKTYETTLK 360  
Db 325 DLPSLAADFVESKDVCKNYAEAKOVFLGMFLYFYARRHPDYSVLLRLAKTYETTLK 384

Qy 361 CAADPHCEYAKVDFEFPKPLVEEPQNLKQCELFQOLGEYKFNALLVRYTKVPQVST 420  
Db 385 CAADPHCEYAKVDFEFPKPLVEEPQNLKQCELFQOLGEYKFNALLVRYTKVPQVST 444

Qy 421 PTLVEVSRLGKVGSKCKKHPKAMPKCAEDYLSVVLNQLCVLHKEKTPVSDRVTKCCTES 480  
Db 445 PTLVEVSRLGKVGSKCKKHPKAMPKCAEDYLSVVLNQLCVLHKEKTPVSDRVTKCCTES 504

Qy 481 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKERQIKQTALVELVHKPKAT 540  
Db 505 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKERQIKQTALVELVHKPKAT 564

Qy 541 KEQLKAVMDDDFAAFVEKCKKADDDKTCFAEEGKKLVAASQAALGL 585  
Db 565 KEQLKAVMDDDFAAFVEKCKKADDDKTCFAEEGKKLVAASQAALGL 609

RESULT 10  
US-09-949-016-11170  
; Sequence 11170, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11170  
; LENGTH: 622  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-11170

Query Match 82.0%; Score 3103; DB 4; Length 622;  
Best Local Similarity 100.0%; Pred. No. 4.3e-84;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHFKDGLGEENFKALVLI AFAQYLQQCPFFEDHVKLVNEVTEFAKTCVADSEAE 60  
Db 38 DAHSEVAHFKDGLGEENFKALVLI AFAQYLQQCPFFEDHVKLVNEVTEFAKTCVADSEAE 97

Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
Db 98 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 157

Qy 121 DVMCTAFHDNEETFLKKLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQADKAACLLP 180  
Db 158 DVMCTAFHDNEETFLKKLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQADKAACLLP 217

Qy 181 KLDELDEGKASSAKORLKASLQKFGERAFAKAWAVARLSORFPKAFPAEVS KLVTDLTK 240  
Db 218 KLDELDEGKASSAKORLKASLQKFGERAFAKAWAVARLSORFPKAFPAEVS KLVTDLTK 277

Qy 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 300  
Db 278 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 337

Qy 301 DLPSLAADFVESKDVCKNYAEAKOVFLGMFLYFYARRHPDYSVLLRLAKTYETTLK 360  
Db 338 DLPSLAADFVESKDVCKNYAEAKOVFLGMFLYFYARRHPDYSVLLRLAKTYETTLK 397

Qy 361 CAADPHCEYAKVDFEFPKPLVEEPQNLKQCELFQOLGEYKFNALLVRYTKVPQVST 420  
Db 398 CAADPHCEYAKVDFEFPKPLVEEPQNLKQCELFQOLGEYKFNALLVRYTKVPQVST 457

Qy 421 PTLVEVSRLGKVGSKCKKHPKAMPKCAEDYLSVVLNQLCVLHKEKTPVSDRVTKCCTES 480  
Db 458 PTLVEVSRLGKVGSKCKKHPKAMPKCAEDYLSVVLNQLCVLHKEKTPVSDRVTKCCTES 517

Qy 481 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKERQIKQTALVELVHKPKAT 540  
Db 518 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKERQIKQTALVELVHKPKAT 577

Qy 541 KEQLKAVMDDDFAAFVEKCKKADDDKTCFAEEGKKLVAASQAALGL 585  
Db 578 KEQLKAVMDDDFAAFVEKCKKADDDKTCFAEEGKKLVAASQAALGL 622

RESULT 11  
US-08-256-938-4  
; Sequence 4, Application US/08256938  
; Patent No. 5685863  
; GENERAL INFORMATION:  
; APPLICANT: Yeh, Patrice  
; TITLE OF INVENTION: NEW POLYPEPTIDES HAVING GRANULOCYTE  
; TITLE OF INVENTION: COLONY STIMULATING ACTIVITY, PREPARATION THEREOF AND  
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING SAID POLYPEPTIDES  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Rhone-Poulenc Rorer Inc.  
; STREET: 500 Arcola Road, 3C43  
; CITY: Collegeville  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19426  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Macintosh  
; OPERATING SYSTEM: System 7.1  
; SOFTWARE: Word 5.0 (PatentIn)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/256,938  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR 92/01065  
; FILING DATE: 31-JAN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Goodman, Rosanne  
; REGISTRATION NUMBER: 32,534  
; REFERENCE/DOCKET NUMBER: ST92007-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (610) 454-3817  
; TELEFAX: (610) 454-3808  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 787 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-256-938-4  
Query Match 82.0%; Score 3103; DB 1; Length 787;  
Best Local Similarity 100.0%; Pred. No. 6.5e-84;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DAHSEVAHFRKDLGFENFALVLIAPAOYLQCPFDHVKLVNEVTEPAKTCVADESAE 60  
DB 203 DAHSEVAHFRKDLGFENFALVLIAPAOYLQCPFDHVKLVNEVTEPAKTCVADESAE 262  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFIQLHKDDNPNLRLVRPEV 120  
DB 263 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFIQLHKDDNPNLRLVRPEV 322  
QY 121 DVNCTAFHNDNEETFLKYLVEIARRHPYFYVAPELLFPAKYKAAFTCCOAAADKAACLLP 180  
DB 323 DVNCTAFHNDNEETFLKYLVEIARRHPYFYVAPELLFPAKYKAAFTCCOAAADKAACLLP 382  
QY 181 KLDELDEGKASAKORLKCASQKGERAFKAWAVARLSQRPKAEFVSKLVTDLTK 240  
DB 383 KLDELDEGKASAKORLKCASQKGERAFKAWAVARLSQRPKAEFVSKLVTDLTK 442  
QY 241 VHTTECHGDLLECDRADRLAKYICENQDSISSKLEKCECKPLLEKSHCIAEYVNDMPA 300  
DB 443 VHTTECHGDLLECDRADRLAKYICENQDSISSKLEKCECKPLLEKSHCIAEYVNDMPA 502  
QY 301 DLPSLAADFVESKDVCKNVAEAKDVLGMPLYEYARRHPDYSVVLRLAKTYETTLEKC 360  
DB 503 DLPSLAADFVESKDVCKNVAEAKDVLGMPLYEYARRHPDYSVVLRLAKTYETTLEKC 562

QY 361 CAAADPHECYAKVFDEFKELVEEPQNLIKONCELFQQLGEYKFPQNALLVRYTKKVPQVST 420  
DB 563 CAAADPHECYAKVFDEFKELVEEPQNLIKONCELFQQLGEYKFPQNALLVRYTKKVPQVST 622  
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTCKCTES 480  
DB 623 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTCKCTES 682  
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVKKHPKAT 540  
DB 683 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVKKHPKAT 742  
QY 541 KEQLKAVMDDDFAAFVEKCKCKADDKETCFABEGKGLVAASQAALGL 585  
DB 743 KEQLKAVMDDDFAAFVEKCKCKADDKETCFABEGKGLVAASQAALGL 787  
RESULT 12  
US-08-797-689-16  
; Sequence 16, Application US/08797689  
; Patent No. 5876969  
; GENERAL INFORMATION:  
; APPLICANT: Fleer, Reinhard  
; APPLICANT: Fournier, Alain  
; APPLICANT: Guittou, Jean-Dominique  
; APPLICANT: Jung, Gerard  
; APPLICANT: Yeh, Patrice  
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES, COMPOSITION  
; TITLE OF INVENTION: PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Rhone-Poulenc Rorer Inc.  
; STREET: 500 Arcola Road, 3C43  
; CITY: Collegeville  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19426  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Macintosh  
; OPERATING SYSTEM: System 7.1  
; SOFTWARE: Word 5.1 (PatentIn)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/797,689  
; FILING DATE: 31-JAN-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/256,927  
; FILING DATE: 28-JUL-1994  
; APPLICATION NUMBER: FR 92/01064  
; FILING DATE: 31-JAN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/FR93/00085  
; FILING DATE: 28-JAN-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith Ph.D. Julie K.  
; REGISTRATION NUMBER: P-36,619  
; REFERENCE/DOCKET NUMBER: ST92006-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (610) 454-3839  
; TELEFAX: (610) 454-3808  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 787 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-797-689-16

Query Match 82.0%; Score 3103; DB 2; Length 787;  
Best Local Similarity 100.0%; Pred. No. 6.5e-84;

Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHFKDGLGEENFKALVLIAPAQYLQCCPFEDHVKLVNVEVTEFAKTCVADSEAE 60  
Db 203 DAHSEVAHFKDGLGEENFKALVLIAPAQYLQCCPFEDHVKLVNVEVTEFAKTCVADSEAE 262

Qy 61 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
Db 263 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 322

Qy 121 DVMTAFHDNEETFLKKYLYEIARRHPYFYVAPELLFFAKRYKAAFTCCQAAADKAACLLP 180  
Db 323 DVMTAFHDNEETFLKKYLYEIARRHPYFYVAPELLFFAKRYKAAFTCCQAAADKAACLLP 382

Qy 181 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSQRFPAKFAEVSCLVTDLT 240  
Db 383 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSQRFPAKFAEVSCLVTDLT 442

Qy 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLECCCKPLLEKSHCIAEVENDEMPA 300  
Db 443 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLECCCKPLLEKSHCIAEVENDEMPA 502

Qy 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVVLLRLAKTYETTTLEKC 360  
Db 503 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVVLLRLAKTYETTTLEKC 562

Qy 361 CAAADPHECVAKVDFEKFPLVEEPQNLIKONCELFQELGKVFQNALLVRTKKVQVST 420  
Db 563 CAAADPHECVAKVDFEKFPLVEEPQNLIKONCELFQELGKVFQNALLVRTKKVQVST 622

Qy 421 PTLVEVSRLNGKVGSKCKGHPKAPCAEDYLSVNLNQLCVLHEKTPVSDRVTCKCTES 480  
Db 623 PTLVEVSRLNGKVGSKCKGHPKAPCAEDYLSVNLNQLCVLHEKTPVSDRVTCKCTES 682

Qy 481 LVNRRPCFSALVDETYVPKFEFNAETFTFHADICTLSEKERQIKKQATLVELVXHKPKAT 540  
Db 683 LVNRRPCFSALVDETYVPKFEFNAETFTFHADICTLSEKERQIKKQATLVELVXHKPKAT 742

Qy 541 KEQLKAVMDDFAAFVEKCKCADDKTCFAEBGKKLVAASQAALGL 585  
Db 743 KEQLKAVMDDFAAFVEKCKCADDKTCFAEBGKKLVAASQAALGL 787

RESULT 13  
US-09-984-186-16  
; Sequence 16, Application US/09984186  
; Patent No. 6686179  
; GENERAL INFORMATION:  
; APPLICANT: Fleer, Reinhard  
; Fournier, Alain  
; Guittion, Jean-Dominique  
; Jung, Gerard  
; Yeh, Patrice  
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,  
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION  
; CONTAINING SAID POLYPEPTIDES  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Rhone-Poulenc Rorer Inc.  
; STREET: 500 Arcola Road, 3C43  
; CITY: Collegeville  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19426  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Macintosh  
; OPERATING SYSTEM: System 7.1  
; SOFTWARE: Word 5.1 (PatentIn)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/984,186  
; FILING DATE: 29-Oct-2001  
; CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/797,689  
; FILING DATE: 31-JAN-1997  
; APPLICATION NUMBER: US 08/256,927  
; FILING DATE: 28-JUL-1994  
; APPLICATION NUMBER: FR 92/01064  
; FILING DATE: 31-JAN-1992  
; APPLICATION NUMBER: PCT/FR93/00085  
; FILING DATE: 28-JAN-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith Ph.D., Julie K.  
; REGISTRATION NUMBER: P-38,619  
; REFERENCE/DOCKET NUMBER: ST92006-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (610) 454-3839  
; TELEFAX: (610) 454-3808  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 787 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 16:  
US-09-984-186-16

Query Match 82.0%; Score 3103; DB 4; Length 787;  
Best Local Similarity 100.0%; Pred. No. 6.5e-84;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHFKDGLGEENFKALVLIAPAQYLQCCPFEDHVKLVNVEVTEFAKTCVADSEAE 60  
Db 203 DAHSEVAHFKDGLGEENFKALVLIAPAQYLQCCPFEDHVKLVNVEVTEFAKTCVADSEAE 262

Qy 61 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
Db 263 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 322

Qy 121 DVMTAFHDNEETFLKKYLYEIARRHPYFYVAPELLFFAKRYKAAFTCCQAAADKAACLLP 180  
Db 323 DVMTAFHDNEETFLKKYLYEIARRHPYFYVAPELLFFAKRYKAAFTCCQAAADKAACLLP 382

Qy 181 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSQRFPAKFAEVSCLVTDLT 240  
Db 383 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSQRFPAKFAEVSCLVTDLT 442

Qy 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLECCCKPLLEKSHCIAEVENDEMPA 300  
Db 443 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLECCCKPLLEKSHCIAEVENDEMPA 502

Qy 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVVLLRLAKTYETTTLEKC 360  
Db 503 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVVLLRLAKTYETTTLEKC 562

Qy 361 CAAADPHECVAKVDFEKFPLVEEPQNLIKONCELFQELGKVFQNALLVRTKKVQVST 420  
Db 563 CAAADPHECVAKVDFEKFPLVEEPQNLIKONCELFQELGKVFQNALLVRTKKVQVST 622

Qy 421 PTLVEVSRLNGKVGSKCKGHPKAPCAEDYLSVNLNQLCVLHEKTPVSDRVTCKCTES 480  
Db 623 PTLVEVSRLNGKVGSKCKGHPKAPCAEDYLSVNLNQLCVLHEKTPVSDRVTCKCTES 682

Qy 481 LVNRRPCFSALVDETYVPKFEFNAETFTFHADICTLSEKERQIKKQATLVELVXHKPKAT 540  
Db 683 LVNRRPCFSALVDETYVPKFEFNAETFTFHADICTLSEKERQIKKQATLVELVXHKPKAT 742

Qy 541 KEQLKAVMDDFAAFVEKCKCADDKTCFAEBGKKLVAASQAALGL 585  
Db 743 KEQLKAVMDDFAAFVEKCKCADDKTCFAEBGKKLVAASQAALGL 787

RESULT 14  
US-08-222-619-3  
; Sequence 3, Application US/08222619

Patent No. 5652352  
GENERAL INFORMATION:  
APPLICANT: Lichenstein, Henri  
APPLICANT: Lyons, David  
APPLICANT: Wurfel, Mark  
APPLICANT: Wright, Samuel  
TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like  
NUMBER OF SEQUENCES: 33  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Amgen Center, Patent Operations/RRC  
STREET: 1840 DeHavilland Drive  
CITY: Thousand Oaks  
STATE: California  
COUNTRY: U.S.  
ZIP: 91320-1789  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/222,619  
FILING DATE:  
CLASSIFICATION: 435  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 609 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-222-619-3

Query Match 81.9%; Score 3099; DB 1; Length 609;  
Best Local Similarity 99.8%; Pred. No. 5.5e-84;  
Matches 584; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY	1	DAKSEVAHRFKDLGGENFKALVLIAPAOYLOQCPEDHVKLVNEVTEFAKTCVADES	60
DB	25	DAKSEVAHRFKDLGGENFKALVLIAPAOYLOQCPEDHVKLVNEVTEFAKTCVADES	84
QY	61	NCDSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV	120
DB	85	NCDSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV	144
QY	121	DVNCCTAFHDNEETFLKKLYEIAARRHPYFYAPELLFFAKRYKAAFTCCOAAKACLLP	180
DB	145	DVNCCTAFHDNEETFLKKLYEIAARRHPYFYAPELLFFAKRYKAAFTCCOAAKACLLP	204
QY	181	KLDELDEGKASSAKORLKCASLOKTFGERAFKAWAVARLSQRPFAEFVSKLVTDLTK	240
DB	205	KLDELDEGKASSAKORLKCASLOKTFGERAFKAWAVARLSQRPFAEFVSKLVTDLTK	264
QY	241	VHTECCGHDLLCADDRAADLAKYICENQDISSKLKECCCKPILLESKSHCIAEVENDEMPA	300
DB	265	VHTECCGHDLLCADDRAADLAKYICENQDISSKLKECCCKPILLESKSHCIAEVENDEMPA	324
QY	301	DLPSLAADFVESKDVCNKYAEAKDVLGMFLFYIARRHPDYSVWLLRLAKTYETTLK	360
DB	325	DLPSLAADFVESKDVCNKYAEAKDVLGMFLFYIARRHPDYSVWLLRLAKTYETTLK	384
QY	361	CAAADPHECYAKVDFEKPPIVEEPQNLIKONCELPQOLGEYKQNALLVRYTKVPQVST	420
DB	385	CAAADPHECYAKVDFEKPPIVEEPQNLIKONCELPQOLGEYKQNALLVRYTKVPQVST	444
QY	421	PTLVEVSRNLGKVGSKCKCKPCKAKRMPCAEDYLSVNLNQLCVLHEKTPVSDRVTKCCTES	480
DB	445	PTLVEVSRNLGKVGSKCKCKPCKAKRMPCAEDYLSVNLNQLCVLHEKTPVSDRVTKCCTES	504
QY	481	LVNRRPCFSALEVDVETVVPKFEFNAETFTFHADICTISEKERQIKKQATALVELVKHPKAT	540
DB	505	LVNRRPCFSALEVDVETVVPKFEFNAETFTFHADICTISEKERQIKKQATALVELVKHPKAT	564

QY 541 KEOLKAVMDDFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 585  
DB 565 KEOLKAVMDDFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 609  
RESULT 15  
US-08-433-037-4  
Sequence 4, Application US/08433037  
Patent No. 5707828  
GENERAL INFORMATION:  
APPLICANT: Sreekrishna, Kotikanyadan  
APPLICANT: Barr, Kathryn A.  
APPLICANT: Brierley, Russell A.  
APPLICANT: Thill, Gregory P.  
APPLICANT: Tschopp, Jueg P.  
TITLE OF INVENTION: EXPRESSION OF HUMAN SERUM ALBUMIN IN  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Scully, Scott, Murphy & Presser  
STREET: 400 Garden City Plaza  
CITY: Garden City  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 11530-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/433,037  
FILING DATE: 03-MAY-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Digiglio, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 9108Z  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 742-4343  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 609 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-433-037-4

Query Match 81.9%; Score 3099; DB 1; Length 609;  
Best Local Similarity 99.8%; Pred. No. 5.5e-84;  
Matches 584; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	DAKSEVAHRFKDLGGENFKALVLIAPAOYLOQCPEDHVKLVNEVTEFAKTCVADES	60
DB	25	DAKSEVAHRFKDLGGENFKALVLIAPAOYLOQCPEDHVKLVNEVTEFAKTCVADES	84
QY	61	NCDSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV	120
DB	85	NCDSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV	144
QY	121	DVNCCTAFHDNEETFLKKLYEIAARRHPYFYAPELLFFAKRYKAAFTCCOAAKACLLP	180
DB	145	DVNCCTAFHDNEETFLKKLYEIAARRHPYFYAPELLFFAKRYKAAFTCCOAAKACLLP	204
QY	181	KLDELDEGKASSAKORLKCASLOKTFGERAFKAWAVARLSQRPFAEFVSKLVTDLTK	240
DB	205	KLDELDEGKASSAKORLKCASLOKTFGERAFKAWAVARLSQRPFAEFVSKLVTDLTK	264
QY	241	VHTECCGHDLLCADDRAADLAKYICENQDISSKLKECCCKPILLESKSHCIAEVENDEMPA	300

Db 265 VHTCCGDLLECCADDRADLAKYICENODSISSKLKECCCKPLLEKSHCIAEVENDEMPA 324  
Qy 301 DLPSLAADFVSKYAKVDEPKPLVEEPONLIKONCELFQGLYVARRHPDYSVLLLRRLAKTYETTLK 360  
Db 325 DLPSLAADFVSKYAKVDEPKPLVEEPONLIKONCELFQGLYVARRHPDYSVLLLRRLAKTYETTLK 384  
Qy 361 CAAADPHCYAKVDEPKPLVEEPONLIKONCELFQGLYVARRHPDYSVLLLRRLAKTYETTLK 420  
Db 385 CAAADPHCYAKVDEPKPLVEEPONLIKONCELFQGLYVARRHPDYSVLLLRRLAKTYETTLK 444  
Qy 421 PTLVEVSRLGKVGSKCKHPEAKMPCABDYLSVNLQCLVLEKTPVSDRVTKCCTES 480  
Db 445 PTLVEVSRLGKVGSKCKHPEAKMPCABDYLSVNLQCLVLEKTPVSDRVTKCCTES 504  
Qy 481 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVKKPKAT 540  
Db 505 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVKKPKAT 564  
Qy 541 KEQLKAVMDDFAAFVEKCKCCKADDDKTCFAEKGKLVAAASQAALGL 585  
Db 565 KEQLKAVMDDFAAFVEKCKCCKADDDKTCFAEKGKLVAAASQAALGL 609

RESULT 16  
US-08-897-956A-2  
; Sequence 2, Application US/08897956A  
; Patent No. 6423512  
; GENERAL INFORMATION:  
; APPLICANT: Mary Ellen Digan  
; APPLICANT: Philip Lake  
; APPLICANT: Hermann Gram  
; TITLE OF INVENTION: Fusion Polypeptides  
; FILE REFERENCE: 600-7244/CPA  
; CURRENT APPLICATION NUMBER: US/08/897,956A  
; CURRENT FILING DATE: 1997-07-21  
; PRIOR APPLICATION NUMBER: 60/022,689  
; PRIOR FILING DATE: 1996-07-26  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 609  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-08-897-956A-2

Query Match 81.9%; Score 3099; DB 4; Length 609;  
Best Local Similarity 99.8%; Pred. No. 5.5e-84;  
Matches 584; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
Db 25 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 84

Qy 61 NCDKSLHTLFGDKLTATLRETYGEMADCCAKQEPERNECFQHKDNDPNLPLVRPEV 120  
Db 85 NCDKSLHTLFGDKLTATLRETYGEMADCCAKQEPERNECFQHKDNDPNLPLVRPEV 144

Qy 121 DVMCTAFHNDNETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180  
Db 145 DVMCTAFHNDNETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 204

Qy 181 KLDELDRGKASSAKQRLKCSLOKFGGERAFKAWAVARLSORFPKAEFAEVSCLVTLTK 240  
Db 205 KLDELDRGKASSAKQRLKCSLOKFGGERAFKAWAVARLSORFPKAEFAEVSCLVTLTK 264

Qy 241 VHTCCGDLLECCADDRADLAKYICENODSISSKLKECCCKPLLEKSHCIAEVENDEMPA 300  
Db 265 VHTCCGDLLECCADDRADLAKYICENODSISSKLKECCCKPLLEKSHCIAEVENDEMPA 324

Qy 301 DLPSLAADFVSKYAKVDEPKPLVEEPONLIKONCELFQGLYVARRHPDYSVLLLRRLAKTYETTLK 360  
Db 325 DLPSLAADFVSKYAKVDEPKPLVEEPONLIKONCELFQGLYVARRHPDYSVLLLRRLAKTYETTLK 384

Qy 361 CAAADPHCYAKVDEPKPLVEEPONLIKONCELFQGLYVARRHPDYSVLLLRRLAKTYETTLK 420  
Db 385 CAAADPHCYAKVDEPKPLVEEPONLIKONCELFQGLYVARRHPDYSVLLLRRLAKTYETTLK 444  
Qy 421 PTLVEVSRLGKVGSKCKHPEAKMPCABDYLSVNLQCLVLEKTPVSDRVTKCCTES 480  
Db 445 PTLVEVSRLGKVGSKCKHPEAKMPCABDYLSVNLQCLVLEKTPVSDRVTKCCTES 504  
Qy 481 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVKKPKAT 540  
Db 505 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVKKPKAT 564  
Qy 541 KEQLKAVMDDFAAFVEKCKCCKADDDKTCFAEKGKLVAAASQAALGL 585  
Db 565 KEQLKAVMDDFAAFVEKCKCCKADDDKTCFAEKGKLVAAASQAALGL 609

RESULT 17  
PCT-US95-04075-3  
; Sequence 3, Application PC/TUS9504075  
; GENERAL INFORMATION:  
; APPLICANT: AMGEN INC.  
; TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like  
; TITLE OF INVENTION: Protein  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Amgen Center, Patent Operations/RRC  
; STREET: 1840 DeHavilland Drive  
; CITY: Thousand Oaks  
; STATE: California  
; COUNTRY: U.S.  
; ZIP: 91320-1789  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/04075  
; FILING DATE:  
; CLASSIFICATION:  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 609 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
PCT-US95-04075-3

Query Match 81.9%; Score 3099; DB 5; Length 609;  
Best Local Similarity 99.8%; Pred. No. 5.5e-84;  
Matches 584; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
Db 25 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 84

Qy 61 NCDKSLHTLFGDKLTATLRETYGEMADCCAKQEPERNECFQHKDNDPNLPLVRPEV 120  
Db 85 NCDKSLHTLFGDKLTATLRETYGEMADCCAKQEPERNECFQHKDNDPNLPLVRPEV 144

Qy 121 DVMCTAFHNDNETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180  
Db 145 DVMCTAFHNDNETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 204

Qy 181 KLDELDRGKASSAKQRLKCSLOKFGGERAFKAWAVARLSORFPKAEFAEVSCLVTLTK 240  
Db 205 KLDELDRGKASSAKQRLKCSLOKFGGERAFKAWAVARLSORFPKAEFAEVSCLVTLTK 264

Qy 241 VHTCCGDLLECCADDRADLAKYICENODSISSKLKECCCKPLLEKSHCIAEVENDEMPA 300  
Db 265 VHTCCGDLLECCADDRADLAKYICENODSISSKLKECCCKPLLEKSHCIAEVENDEMPA 324



```
QY 301 DLPSLAADFVESKDVCKNVAEKDVFLGMFLYFYARRHPDYSVVLRLRLAKTYETTTLEKC 360
DB 325 DLPSLAADFVESKDVCKNVAEKDVFLGMFLYFYARRHPDYSVVLRLRLAKTYETTTLEKC 384
QY 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCBLFEQLGKYKFNALLVRYTKKVPQVST 420
DB 385 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCBLFEQLGKYKFNALLVRYTKKVPQVST 444
QY 421 PTLVEVSRLNGKVGSKCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 445 PTLVEVSRLNGKVGSKCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDYTVPKFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 540
DB 505 LVNRRPCFSALEVDYTVPKFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 564
QY 541 KEOLKAVMDDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 585
DB 565 KEOLKAVMDDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 609

RESULT 18
US-08-448-196A-3
; Sequence 3, Application US/08448196A
; Patent No. 5780594
; GENERAL INFORMATION:
; APPLICANT: CARTER, DANIEL C.
; TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS
; TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR
; TITLE OF INVENTION: RELATED PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NASA
; STREET: MARSHALL SPACE FLIGHT CENTER
; CITY: HUNTSVILLE
; STATE: ALABAMA
; COUNTRY: USA
; ZIP: 35812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448,196A
; FILING DATE: 23-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: BROAD JR., ROBERT L.
; REGISTRATION NUMBER: 18,757
; REFERENCE/DOCKET NUMBER: XX/MFS-28402-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 205-544-0021
; TELEFAX: 205-544-0258
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
US-08-448-196A-3

Query Match 81.7%; Score 3093; DB 1; Length 585;
Best Local Similarity 99.7%; Pred. No. 7.7e-84;
Matches 583; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DAHKSVAHRFKDLGLENFKALVLIAPAQYLQCCPFEDHVKLVNEVTEFAKTCVADSSAE 60
DB 1 DAHKSVAHRFKDLGLENFKALVLIAPAQYLQCCPFEDHVKLVNEVTEFAKTCVADSSAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPRPEV 120
QY 121 DVNCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP 180
DB 121 DVNCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP 180
QY 181 KLDELDEGKASAKORLAKCASLQKFGERAFAKAWAVARLSQRPPKAEFAEVSKLVDLTG 240
DB 181 KLDELDEGKASAKORLAKCASLQKFGERAFAKAWAVARLSQRPPKAEFAEVSKLVDLTG 240
QY 241 VHTTECHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEYENDEMPA 300
DB 241 VHTTECHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEYENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNVAEKDVFLGMFLYFYARRHPDYSVVLRLRLAKTYETTTLEKC 360
DB 301 DLPSLAADFVESKDVCKNVAEKDVFLGMFLYFYARRHPDYSVVLRLRLAKTYETTTLEKC 360
QY 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCBLFEQLGKYKFNALLVRYTKKVPQVST 420
DB 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCBLFEQLGKYKFNALLVRYTKKVPQVST 420
QY 421 PTLVEVSRLNGKVGSKCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRLNGKVGSKCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDYTVPKFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 540
DB 481 LVNRRPCFSALEVDYTVPKFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 540
QY 541 KEOLKAVMDDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 585
DB 541 KEOLKAVMDDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 585

RESULT 19
US-08-984-176-1
; Sequence 1, Application US/08984176
; Patent No. 5948609
; GENERAL INFORMATION:
; APPLICANT: CARTER, DANIEL C
; APPLICANT: HO, JOSEPH X
; APPLICANT: RUKER, FLORIAN
; TITLE OF INVENTION: OXYGEN-TRANSPORTING ALBUMIN-BASED BLOOD REPLACEMENT
; FILE REFERENCE: 08/984,176
; CURRENT APPLICATION NUMBER: US/08/984,176
; CURRENT FILING DATE: 1997-12-03
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-984-176-1

Query Match 81.7%; Score 3093; DB 2; Length 585;
Best Local Similarity 99.7%; Pred. No. 7.7e-84;
Matches 583; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DAHKSVAHRFKDLGLENFKALVLIAPAQYLQCCPFEDHVKLVNEVTEFAKTCVADSSAE 60
DB 1 DAHKSVAHRFKDLGLENFKALVLIAPAQYLQCCPFEDHVKLVNEVTEFAKTCVADSSAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPRPEV 120
QY 121 DVNCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP 180
DB 121 DVNCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP 180
```

Db 121 DVMCTAFHDNEETFLKYLVEIARRHPYFYFAPELLFFAKRYKAAFTCCQADKAAACLLP 180  
Qy 181 KLDELDEGKASSAKQRLKASLOKFGGERAFKAWAVARLSQRFPAKFAEVSKLVTDLTK 240  
Db 181 KLDELDEGKASSAKQRLKASLOKFGGERAFKAWAVARLSQRFPAKFAEVSKLVTDLTK 240  
Qy 241 VHTCCGHDLLLEACADDRADLAKYICENQDSISSKLECCCKPDLLEKSHCIAEVENDEMPA 300  
Db 241 VHTCCGHDLLLEACADDRADLAKYICENQDSISSKLECCCKPDLLEKSHCIAEVENDEMPA 300  
Qy 301 DLPSLAADFVESKDVCKNYABAKDVFLGMFLYEVARRHPDYSVVLLRLAKTYETTLK 360  
Db 301 DLPSLAADFVESKDVCKNYABAKDVFLGMFLYEVARRHPDYSVVLLRLAKTYETTLK 360  
Qy 361 CAAADPHCEYAKVDFEFPKPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKVPQVST 420  
Db 361 CAAADPHCEYAKVDFEFPKPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKVPQVST 420  
Qy 421 PTLVEVRNLGKVGSKCKHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
Db 421 PTLVEVRNLGKVGSKCKHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
Qy 481 LVNRRPCFSALEVDVETVYKPFNAETFTFHADICTLSEKEROIKKQATLVELVHKPKAT 540  
Db 481 LVNRRPCFSALEVDVETVYKPFNAETFTFHADICTLSEKEROIKKQATLVELVHKPKAT 540  
Qy 541 KEQLKAVMDDFAAFVEKCKADDDKTCFAEKGKLVAAASQAALGL 585  
Db 541 KEQLKAVMDDFAAFVEKCKADDDKTCFAEKGKLVAAASQAALGL 585

RESULT 20  
US-08-448-196A-5  
; Sequence 5, Application US/08448196A  
; Patent No. 5780594  
; GENERAL INFORMATION:  
; APPLICANT: CARTER, DANIEL C.  
; TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS  
; TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR  
; TITLE OF INVENTION: RELATED PROTEINS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NASA  
; STREET: MARSHALL SPACE FLIGHT CENTER  
; CITY: HUNTSVILLE  
; STATE: ALABAMA  
; COUNTRY: USA  
; ZIP: 35812  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/448,196A  
; FILING DATE: 23-MAY-1995  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BROAD JR., ROBERT L.  
; REGISTRATION NUMBER: 18,757  
; REFERENCE/DOCKET NUMBER: XX/MFS-28402-2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 205-544-0021  
; TELEFAX: 205-544-0258  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 583 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: N-terminal

US-08-448-196A-5  
Query Match 65.0%; Score 2458.9; DB 1; Length 583;  
Best Local Similarity 75.8%; Pred. No. 4.6e-65;  
Matches 442; Conservative 70; Mismatches 70; Indels 1; Gaps 1;  
Qy 1 DAKSEVAHFKDLGEENFKALVLIAPAOYLQOCPPEDHVKLVNVEVTEFAKTCVADSEAE 60  
Db 1 DTHKSEIAHRFNDLGEKHFGLVLAQFQYLQOCPPEDHVKLVNVEVTEFAKTCVADSEAE 60  
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGENADCCAKQEPERNECFLOHKDDNPNLPLVPEV 120  
Db 61 NCDKSLHTLFGDKLCTVATLRETYGENADCCAKQEPERNECFLOHKDDNPNLPLVPEV 120  
Qy 121 DVMCTAFHDNEETFLKYLVEIARRHPYFYFAPELLFFAKRYKAAFTCCQADKAAACLLP 180  
Db 121 DVMCTAFHDNEETFLKYLVEIARRHPYFYFAPELLFFAKRYKAAFTCCQADKAAACLLP 180  
Qy 181 KLDELDEGKASSAKQRLKASLOKFGGERAFKAWAVARLSQRFPAKFAEVSKLVTDLTK 240  
Db 181 KLDELDEGKASSAKQRLKASLOKFGGERAFKAWAVARLSQRFPAKFAEVSKLVTDLTK 240  
Qy 241 VHTCCGHDLLLEACADDRADLAKYICENQDSISSKLECCCKPDLLEKSHCIAEVENDEMPA 300  
Db 241 VHTCCGHDLLLEACADDRADLAKYICENQDSISSKLECCCKPDLLEKSHCIAEVENDEMPA 300  
Qy 301 DLPSLAADFVESKDVCKNYABAKDVFLGMFLYEVARRHPDYSVVLLRLAKTYETTLK 360  
Db 301 DLPSLAADFVESKDVCKNYABAKDVFLGMFLYEVARRHPDYSVVLLRLAKTYETTLK 360  
Qy 361 CAAADPHCEYAKVDFEFPKPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKVPQVST 420  
Db 361 CAAADPHCEYAKVDFEFPKPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKVPQVST 420  
Qy 421 PTLVEVRNLGKVGSKCKHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
Db 421 PTLVEVRNLGKVGSKCKHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
Qy 481 LVNRRPCFSALEVDVETVYKPFNAETFTFHADICTLSEKEROIKKQATLVELVHKPKAT 540  
Db 481 LVNRRPCFSALEVDVETVYKPFNAETFTFHADICTLSEKEROIKKQATLVELVHKPKAT 540  
Qy 541 KEQLKAVMDDFAAFVEKCKADDDKTCFAEKGKLVAAASQAALGL 583  
Db 541 KEQLKAVMDDFAAFVEKCKADDDKTCFAEKGKLVAAASQAALGL 583

RESULT 21  
US-08-448-196A-4  
; Sequence 4, Application US/08448196A  
; Patent No. 5780594  
; GENERAL INFORMATION:  
; APPLICANT: CARTER, DANIEL C.  
; TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS  
; TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR  
; TITLE OF INVENTION: RELATED PROTEINS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NASA  
; STREET: MARSHALL SPACE FLIGHT CENTER  
; CITY: HUNTSVILLE  
; STATE: ALABAMA  
; COUNTRY: USA  
; ZIP: 35812  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/448,196A  
; FILING DATE: 23-MAY-1995  
; CLASSIFICATION: 530

```
; ATTORNEY/AGENT INFORMATION:
; NAME: BROAD JR., ROBERT L.
; REGISTRATION NUMBER: 18,757
; REFERENCE/DOCKET NUMBER: XX/MFS-28402-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 205-544-0021
; TELEFAX: 205-544-0258
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 583 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; US-08-448-196A-4

Query Match 64.8%; Score 2450.9; DB 1; Length 583;
Best Local Similarity 75.8%; Pred. No. 7.9e-65;
Matches 442; Conservative 71; Mismatches 69; Indels 1; Gaps 1;

QY 1 DAHSEVAHRFDLGBENFKALVLIAPAOYLQCCPDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DTHKSEIAHRFDLGBEHFKGLVLIASFQYLQCCPDHVKLVNELTEFAKTCVADESAA 60
QY 61 NCDKSLHTLFGDKLCTVATRETYGEMADCCAKQEPERNECFLOHDDNPNLRLVRPEV 120
DB 61 GCEKSLHTLFGDELCKVSLRETYGEMADCCAKQEPERNECFLOHDDNPNLRLVRPEV 120
QY 121 DVNCTAFHNEETFLKKLYEIAARRHPYFYAPPELLFAKRYKAAFTCCQAADKACLLP 180
DB 120 NTICDFEKADEKKFWGKLYEIVARRHPYFYAPPELLFYANKYNGVFQCCQAEKACLLP 179
QY 181 KLDELDEGKASSAKORLKCASLOKFGERAFAKAWARLSQRPKAEFAEVSKLVTDLTK 240
DB 180 KIETREKVLASSARQRLCASLOKFGERALKAWARLSQRPKAEFAEVSKLVTDLTK 239
QY 241 VHTCCGHDLLCADDRAADLAKYICENQDTSISKLKECCBKPLEKSHCIAEVDKEMPA 300
DB 240 VHKECCHGDLLECCADDRAADLAKYICENQDTSISKLKECCBKPLEKSHCIAEVDKEMPA 299
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLFYARRHPDYSVLLRLAKTYETTLK 360
DB 300 NLPLTADFAEDKDVCKNYQEAADFLGSLFYYSRRHPDYAVSVLLRLAKTYETTLK 359
QY 361 CAADAPHECYAKVDFEKPVLVEPQNLIKQNCLEFQOLGEYKFGQNALIVRYTRKVPQVST 420
DB 360 CAKDHPACYSTVFDKLKHLVDEPQNLIKQNCDOFEKLGEGYFGQNALIVRYTRKVPQVST 419
QY 421 PTLVEVSRLNGKVGSKCKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 420 PTLVEVSRLNGKVGTRCCTKPESERMPCETEDYLSLILNRLCVLHEKTPVSEKVTTCCTES 479
QY 481 LVNRRPCFSALEVDYETVVPKEFNAETFTFHADICTLSEKEROIKQOTALVELVKKPKAT 540
DB 480 LVNRRPCFSALTDETVVVPKAFDEKLTFTFHADICTLPTDEKQIKQOTALVELVKKPKAT 539
QY 541 KEOLKAWMDFAAFVEKCCKADKDETCFABEGKGLVAASQAAL 583
DB 540 BEQLKTWENFVAFVDCCAADDEKACFAVEGPKLVVSTOTAL 582
```

## RESULT 22

```
US-08-448-196A-6
; Sequence 6, Application US/08448196A
; Patent No. 5780594
; GENERAL INFORMATION:
; APPLICANT: CARTER, DANIEL C.
; TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS
; TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR
; TITLE OF INVENTION: RELATED PROTEINS
; NUMBER OF SEQUENCES: 9
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NASA
; STREET: MARSHALL SPACE FLIGHT CENTER
; CITY: HUNTSVILLE
; STATE: ALABAMA
; COUNTRY: USA
; ZIP: 35812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448,196A
; FILING DATE: 23-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: BROAD JR., ROBERT L.
; REGISTRATION NUMBER: 18,757
; REFERENCE/DOCKET NUMBER: XX/MFS-28402-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 205-544-0021
; TELEFAX: 205-544-0258
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 583 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; US-08-448-196A-6
```

```
Query Match 64.3%; Score 2432.9; DB 1; Length 583;
Best Local Similarity 75.0%; Pred. No. 2.7e-64;
Matches 437; Conservative 73; Mismatches 72; Indels 1; Gaps 1;

QY 1 DAHSEVAHRFDLGBENFKALVLIAPAOYLQCCPDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DTHKSEIAHRFDLGBENFKGLVLIASFQYLQCCPDHVKLVNELTEFAKTCVADESAA 60
QY 61 NCDKSLHTLFGDKLCTVATRETYGEMADCCAKQEPERNECFLOHDDNPNLRLVRPEV 120
DB 61 GCEKSLHTLFGDELCKVSLRETYGEMADCCAKQEPERNECFLOHDDNPNLRLVRPEV 120
QY 121 DVNCTAFHNEETFLKKLYEIAARRHPYFYAPPELLFAKRYKAAFTCCQAADKACLLP 180
DB 120 DTLCAEFKADEKKFWGKLYEIVARRHPYFYAPPELLFYANKYNGVFQCCQAEKACLLP 179
QY 181 KLDELDEGKASSAKORLKCASLOKFGERAFAKAWARLSQRPKAEFAEVSKLVTDLTK 240
DB 180 KIDAMREKVLASSARQRLCASLOKFGERALKAWARLSQRPKAEFAEVSKLVTDLTK 239
QY 241 VHTCCGHDLLCADDRAADLAKYICENQDTSISKLKECCBKPLEKSHCIAEVDKEMPA 300
DB 240 VHKECCHGDLLECCADDRAADLAKYICENQDTSISKLKECCBKPLEKSHCIAEVDKEMPA 299
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLFYARRHPDYSVLLRLAKTYETTLK 360
DB 300 NLPLTADFAEDKDVCKNYQEAADFLGSLFYYSRRHPDYAVSVLLRLAKTYETTLK 359
QY 361 CAADAPHECYAKVDFEKPVLVEPQNLIKQNCLEFQOLGEYKFGQNALIVRYTRKVPQVST 420
DB 360 CAKDHPACYSTVFDKLKHLVDEPQNLIKQNCLEFQOLGEYKFGQNALIVRYTRKVPQVST 419
QY 421 PTLVEVSRLNGKVGSKCKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 420 PTLVEVSRLNGKVGTRCCTKPESERMPCETEDYLSLILNRLCVLHEKTPVSEKVTTCCTES 479
QY 481 LVNRRPCFSALEVDYETVVPKEFNAETFTFHADICTLSEKEROIKQOTALVELVKKPKAT 540
DB 480 LVNRRPCFSDLTDETVVVPKAFDEKLTFTFHADICTLPTDEKQIKQOTALVELVKKPKAT 539
```

Qy 541 KEQLKAVMDPFAFVEKCKADDKETCFABEGKGLVAASQAAL 583  
Db 540 DEQLKTWENFVAFVDKCAADDEKGCFLVLEGGPKLVAASQAAL 582

## RESULT 23

US-08-448-196A-7

; Sequence 7, Application US/08448196A

; Patent No. 5780594

; GENERAL INFORMATION:

; APPLICANT: CARTER, DANIEL C.

; TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS

; TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR

; TITLE OF INVENTION: RELATED PROTEINS

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: NASA

; STREET: MARSHALL SPACE FLIGHT CENTER

; CITY: HUNTSVILLE

; STATE: ALABAMA

; COUNTRY: USA

; ZIP: 35812

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible

; SOFTWARE: Patent in Release #1.0, Version #1.30

; CURRENT APPLICATION DATA: US/08/448,196A

; FILING DATE: 23-MAY-1995

; CLASSIFICATION: 530

; ATTORNEY/AGENT INFORMATION:

; NAME: BROAD JR., ROBERT L.

; REGISTRATION NUMBER: 18,757

; REFERENCE/DOCKET NUMBER: XX/MFS-28402-2

; TELEPHONE: 205-544-0021

; TELEFAX: 205-544-0258

; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 584 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHEICAL: NO

; ANTI-SENSE: NO

; FRAGMENT TYPE: N-terminal

; US-08-448-196A-7

Query Match 64.1%; Score 2426; DB 1; Length 584;

Best Local Similarity 73.4%; Pred No. 4.3e-64;

Matches 428; Conservative 82; Mismatches 73; Indels 0; Gaps 0;

Qy 1 DAHSEVAHFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAG 60

Db 1 EAHKSEIAHRFKDLGEQHFGLVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAG 60

Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPENECFLQHKDNDPNLPRLVRPEV 120

Db 61 NCDKSIHTLFGDKLCTVATLRETYGEMADCCAKQBPENECFLQHKDNDPNLPRLVRPEV 120

Qy 121 DVMTAFHNDTEFLKKLYETAREHPYFAPPELLFAKRYKAAFTCCQADKAACLLP 180

Db 121 EAMCTSFQENPTSFGLHYEHVARHPYFAPPELLYAEKNEVUTQCTSDSKAACLLP 180

Qy 181 KLDELDRSGKASSAKQRLKCSLQKFGRAFAMAVARLSORFPFAEFAEYSLVTLDTK 240

Db 181 KLDVAKELAAVAVPQRMKSSQMGFGRAFAMAVARMSQFPFAEFAEITKLATDVTK 240

Qy 241 VHTCECHGDLLECADRDADLAKYICENQDSISSKLUKECEPFLLEKSHCIAEVENDEMPA 300

Db 241 INKECCHGDLLECADRDADLAKYICENQDSISSKLUKECEPFLLEKSHCIAEVENDEMPA 300

Qy 301 DLPSLAADFVESKDVCKNVAEAKDVFLGMFLYEVARRHPDYSVVLLRLAKTYETTLK 360  
Db 301 DLPSIAADFVEDKEVCNVAEAKDVFLGTLFLYEVARRHPDYSVSLRLAKTYETTLK 360  
Qy 361 CAADPHCEYAKVDFEPKPLVEEPQNLKONCELFQOLGEYKFNALLVRYTKVPQVST 420  
Db 361 CAEGDPACYGTVLAEPQPLVEEPQNLKONCELYEKLGEYGFQNAVLRVYTKAPQVST 420  
Qy 421 PTLVEVSRLNGKVGSKCKHPKAEKMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480  
Db 421 PTLVEAARNLGRVGTCTLPEAQRLPCVEDYLSAILNRLCVLHEKTPVSEKVTCCSGS 480  
Qy 481 LVNRRPCFSALEVDVETVPKFNATETTFHADICTLSEKEROIKQTALVELVKKPKAT 540  
Db 481 LVERRPCFSALTVDVETVPKFAETTFHSDICTLPDKEKQIKQTALVELVKKPKAT 540  
Qy 541 KEQLKAVMDPFAFVEKCKADDKETCFABEGKGLVAASQAAL 583  
Db 541 EDQLKTWMDPFAFVDFKCKAADKONCFATEGPNLVARSKAAL 583

## RESULT 24

US-08-134-638-1

; Sequence 1, Application US/08134638

; Patent No. 5473050

; GENERAL INFORMATION:

; APPLICANT: Strand, Frederick T

; TITLE OF INVENTION: Denatured Bovine Serum Albumin Milk

; TITLE OF INVENTION: Products and Method Therefor

; NUMBER OF SEQUENCES: 1

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Frederick T. Strand

; STREET: P.O. Box 64321

; CITY: Phoenix

; STATE: Arizona

; COUNTRY: USA

; ZIP: 85082-4321

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 5.25 inch, 1.2 Mb storage

; OPERATING SYSTEM: MS-DOS 5.0

; SOFTWARE: Wordperfect 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/134,638

; FILING DATE: 10/12/93

; CLASSIFICATION: 530

; PRIOR APPLICATION NUMBER: N/A

; FILING DATE: N/A

; ATTORNEY/AGENT INFORMATION:

; NAME: Weiss, Harry M

; REGISTRATION NUMBER: 19,497

; REFERENCE/DOCKET NUMBER: 1795P1423

; TELEPHONE: (602) 994-8888

; TELEFAX: (602) 947-2663

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 582

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; US-08-134-638-1

Query Match 63.1%; Score 2389.8; DB 1; Length 582;

Best Local Similarity 74.1%; Pred. No. 5.1e-63;

Matches 432; Conservative 75; Mismatches 74; Indels 2; Gaps 2;

Qy 1 DAHSEVAHFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAG 60

Db 1 DTHKSEIAHRFKDLGEHFKGLVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAG 60



QY 4 KSEVAHR-----FKDLGEENFKALVLIATAFYLOQCPCPDHVKLVNEVTEPAKT 52  
Db 17 ESRTLRNEYGIASILDSYQCTAEISLADLATIFFAQVQEAATYKEVSKVWKDALTAIEK 76  
QY 53 CVADESAENCKSLHTLFGDKLCTVATRETYGEMADCCAKQBPNERCEFLQHKDDNP-N 111  
Db 77 PTGDEQSSGCLNQLPAFLBELCHEKEILEKYGH-SDCCSQSSEGRHNCFLAHKKPTPAS 135  
QY 112 LPRLVRPEVDMCTAFHNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAAFTECCQA 171  
Db 136 IPLFQVPEPVTSCAEYEDRETFWNKFIYIARRHPFLYAPTILLMAARYDKIIPSCCKA 195  
QY 172 ADKAAACLLPKLDELDEBKGASSAKQRLKASQKFGERAFKAWAVARLSORFFPKAEFAEV 231  
Db 196 ENAVECFQTKAAATVTKELRESSLLNHACAVMKNFGTRTFOAITVTKLSQKFTKVNTEI 255  
QY 232 SKLVTDLTKVHTECHGDILLECADRADLAKYICENQDSISSKJKECEKPLEKSHCIA 291  
Db 256 QKLVLDVAHVHCCRGDVLDCLODGEKIMSYICSQDQTLNKNITECCKLTTLERGQCII 315  
QY 292 EVENDENPADLPSLAADFVESKVKYAEAKVFLGMEFLYVYARRHPDYVSVLLRLAK 351  
Db 316 HAENDERPEGLSPNLNRFGLDRDQNFSSGEGKNIFLASFVHEYRRHPQLAVSVILRVAK 375  
QY 352 TYETTLKCCAAADPHECYAKVDFEPLVEEPQNLKONCELPQELGEYKFNQALLVRY 411  
Db 376 GYQELLEKCFQTEPLECQDGEELQKYIQESQALAKRSCGLFKLGEYLYQNAFLVAY 435  
QY 412 TKKYPQVSTPTLVBSVRNLGKVGSKCKHPKAKMPCAEEDYLSVVLNQLCVLHKEKTPVSD 471  
Db 436 TKKAPQUTSSSELMAITRKMAATAATCCQSDKLLACGEAGADIIGHLCIRHEMTVPNP 495  
QY 472 RVTCKCTESLVNRRPCFSALEVDYTPKFNATFTFHADICTLSEKERQIKKOTALVE 531  
Db 496 GVGQCCTSSYANRRPCFSSLVVDYTPPAPSDDKFIHKLDCQAQGVALQTMKEFLIN 555  
QY 532 LVKHKPKATKQKAVMDDFAAFEKCKKADDDKTCFAEEGKLVAAASQAALGL 585  
Db 556 LVKQKQITTEQLEAVIADFSLGLEKCCQGEQVCFAEEGQKLISKTRAAALGV 609

RESULT 27  
PCT-US95-04075-4  
; Sequence 4, Application PC/TUS9504075  
; GENERAL INFORMATION:  
; APPLICANT: AMGEN INC.  
; TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like  
; TITLE OF INVENTION: Protein  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Amgen Center, Patent Operations/RRC  
; STREET: 1840 DeHavilland Drive  
; CITY: Thousand Oaks  
; STATE: California  
; COUNTRY: U.S.  
; ZIP: 91320-1789  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/04075  
; FILING DATE:  
; CLASSIFICATION:  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 609 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein

PCT-US95-04075-4  
Query Match 33.1%; Score 1252.7; DB 5; Length 609;  
Best Local Similarity 39.7%; Pred. No. 2.6e-29;  
Matches 236; Conservative 114; Mismatches 231; Indels 13; Gaps 3;  
QY 4 KSEVAHR-----FKDLGEENFKALVLIATAFYLOQCPCPDHVKLVNEVTEPAKT 52  
Db 17 ESRTLRNEYGIASILDSYQCTAEISLADLATIFFAQVQEAATYKEVSKVWKDALTAIEK 76  
QY 53 CVADESAENCKSLHTLFGDKLCTVATRETYGEMADCCAKQBPNERCEFLQHKDDNP-N 111  
Db 77 PTGDEQSSGCLNQLPAFLBELCHEKEILEKYGH-SDCCSQSSEGRHNCFLAHKKPTPAS 135  
QY 112 LPRLVRPEVDMCTAFHNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAAFTECCQA 171  
Db 136 IPLFQVPEPVTSCAEYEDRETFWNKFIYIARRHPFLYAPTILLMAARYDKIIPSCCKA 195  
QY 172 ADKAAACLLPKLDELDEBKGASSAKQRLKASQKFGERAFKAWAVARLSORFFPKAEFAEV 231  
Db 196 ENAVECFQTKAAATVTKELRESSLLNHACAVMKNFGTRTFOAITVTKLSQKFTKVNTEI 255  
QY 232 SKLVTDLTKVHTECHGDILLECADRADLAKYICENQDSISSKJKECEKPLEKSHCIA 291  
Db 256 QKLVLDVAHVHCCRGDVLDCLODGEKIMSYICSQDQTLNKNITECCKLTTLERGQCII 315  
QY 292 EVENDENPADLPSLAADFVESKVKYAEAKVFLGMEFLYVYARRHPDYVSVLLRLAK 351  
Db 316 HAENDERPEGLSPNLNRFGLDRDQNFSSGEGKNIFLASFVHEYRRHPQLAVSVILRVAK 375  
QY 352 TYETTLKCCAAADPHECYAKVDFEPLVEEPQNLKONCELPQELGEYKFNQALLVRY 411  
Db 376 GYQELLEKCFQTEPLECQDGEELQKYIQESQALAKRSCGLFKLGEYLYQNAFLVAY 435  
QY 412 TKKYPQVSTPTLVBSVRNLGKVGSKCKHPKAKMPCAEEDYLSVVLNQLCVLHKEKTPVSD 471  
Db 436 TKKAPQUTSSSELMAITRKMAATAATCCQSDKLLACGEAGADIIGHLCIRHEMTVPNP 495  
QY 472 RVTCKCTESLVNRRPCFSALEVDYTPKFNATFTFHADICTLSEKERQIKKOTALVE 531  
Db 496 GVGQCCTSSYANRRPCFSSLVVDYTPPAPSDDKFIHKLDCQAQGVALQTMKEFLIN 555  
QY 532 LVKHKPKATKQKAVMDDFAAFEKCKKADDDKTCFAEEGKLVAAASQAALGL 585  
Db 556 LVKQKQITTEQLEAVIADFSLGLEKCCQGEQVCFAEEGQKLISKTRAAALGV 609

RESULT 28  
US-09-949-016-11201  
; Sequence 11201, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11201  
; LENGTH: 612  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-11201  
Query Match 33.1%; Score 1252.7; DB 4; Length 612;

```
Best Local Similarity 39.7%; Pred. No. 2.6e-29;
Matches 236; Conservative 114; Mismatches 231; Indels 13; Gaps 3;

QY 4 KSEVAHR-----FKDLGEENFKALVLIARAYQLQCCPFEDHVKLVNVEVTEFAKT 52
Db
20 ESRTLHRNEYGIASILDYSQCTAEISLADLATIFFAQVQOEATYKEVSKNWKDALTAIEK 79
QY 53 CVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQOEPRNECFLOHKDNP-N 111
Db 80 PTGDEOSSGLENQLPAFLBELCHEKEILEKYGH-SDCCSQSEGRHNCFLAHHKTPAS 138
QY 112 LPLVRPEVDVMTAFHDNEETFLKKYLYEIARRHPFYFAPPELLFFAKRYKAAFTCCQA 171
Db 139 IPLFQVPEPVTSCAEYEDRETFMNKFIYEIARRHPFLYAPTILLWAARYDKIIPSCCKA 198
QY 172 ADKAAACLLPKDLRLDEGRKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPFKAEFV 231
Db 199 ENAVECFQTKAAATVTKELRESSLLNQHACAPMKNFGTRTFQAITVTKLSQKFTKVNFT 258
QY 232 SKLVTDLTKVHTCCCHGDLLEACADDRADLAKYICENQDSISSKLEKCECKPILLEKSHCIA 291
Db 259 QKLVLDVAHVHEHCRRGVDLCLQDGEKINSYICSQODTLSNKITTECKLTTLERGCQCI 318
QY 292 EVNDENMPADPLSLAADFVESKDVCKNYABAKOVFLGMFLYFYARRHPDYSVLLRLAK 351
Db 319 HAENDEKPEGLSPNLNRFGLDRDFNQFSGGEKNIFLASFVHEYSSRRHPQLAVSVILRV 378
QY 352 TYETITLKKCAADPHCEYAKVDEFPKPLVEEPQNLKONCELFEQLEGYKFFQNALIVRY 411
Db 379 GYOELLEKCFQTNPLECQDKGEBELQKYIQESQALAKRSCGLFQKLGYYLQNAFLVAY 438
QY 412 TKKVPQVSTPTLVESVNRNLGKVGSKCKCHPEAKRMPCAEYLSVVLNQLCVLHEKTPVSD 471
Db 439 TKKAPQITSELMATIRKWAATAATCCQLSEDKLLACGEGAADIIGHLCIRHEMTVPNP 498
QY 472 RVTKKCTESLVNRRPCFSALEVDETVYPKEFNAETFTFHADICTLSEKERQIKKOTALVE 531
Db 499 GVGQCCTSSYANRRPCFSSLVVDVETVYPPAFSDDKFIHFKDLCAQAGVALQTMKQEBFLN 558
QY 532 LVKHGKPKATKEQLKAVMDDPFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 585
Db 559 LVKQKQPITTEOELEAVIADFSGLLEKCCQGEQEVCFABEGQKLSKTRAAALGV 612

RESULT 29
US-09-186-949A-2
; Sequence 2, Application US/09186949A
; Patent No. 6416734
; GENERAL INFORMATION:
; APPLICANT: Murgita, Robert A.
; TITLE OF INVENTION: Recombinant Alpha-Fetoprotein For
; TITLE OF INVENTION: Treating and Diagnosing Cancers
; FILE REFERENCE: 06727/004002
; CURRENT APPLICATION NUMBER: US/09/186,949A
; PRIOR FILING DATE: 1998-11-05
; PRIOR FILING DATE: US 08/758,757
; PRIOR FILING DATE: 1996-12-03
; PRIOR APPLICATION NUMBER: US 08/377,311
; PRIOR FILING DATE: 1995-01-24
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-186-949A-2

Query Match 32.0%; Score 1211.3; DB 4; Length 609;
Best Local Similarity 38.8%; Pred. No. 4.3e-28;
Matches 231; Conservative 113; Mismatches 235; Indels 17; Gaps 4;

QY 4 KSEVAHR-----FKDLGEENFKALVLIARAYQLQCCPFEDHVKLVNVEVTEFAKT 52
Db

Query Match 32.0%; Score 1210.9; DB 2; Length 590;
Best Local Similarity 38.9%; Pred. No. 4.2e-28;
Matches 230; Conservative 115; Mismatches 235; Indels 11; Gaps 4;

QY 3 HKSE-----VAHRFKDLGLENFKALVLIAPAOYLQCCPFEDHVKLVNVEVTEFAKTCAVE 57
Db 3 HRNEYGIASILDYSQCTAEISLADLATIFFAQVQOEATYKEVSKNWKDALTAIEKPTGDE 62
QY 58 SAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQOEPRNECFLOHKDNPN---LPR 114
Db 63 QSSGLENQLPAFLBELCHEKEILEKYGH-SDCCSQSEGRHNCFLAHHKTPAAWIPL 119
QY 115 LVRPEVDVMTAFHDNEETFLKKYLYEIARRHPFYFAPPELLFFAKRYKAAFTCCCOADK 174
Db 120 FQVPEPVTSCAEYEDRETFMNKFIYEIARRHPFLYAPTILLWAARYDKIIPSCCKAENA 179
```



Qy	175	AACLLPKLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSQRPFKAEFAEVSKL	234
Db	180	VECFQTKAAATVTKELRESSLLNQHACPMKNFGTRTFQAITVTKLSOKFTKVNFTQKL	239
Qy	235	VTDLTKVHTCCCHGDLLECADDDRADLAKYICENQDSISSKKECCEKPLEKSHCIAEVE	294
Db	240	VLDVAHVHEHCCRADVLDCLQDGEKIMSYICSQODTILSNKITECCKLTTLERGQCIIHAE	299
Qy	295	NDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMPLYEYARRHPDYSVVLLRLAKTYE	354
Db	300	NDEXPEGLSPNLNRPFLGDRDFNQFSSGCKNIFLASFVHEYSRHPQLAVSVILEVAKGYQ	359
Qy	355	TTLKCCAAADPHCEYAKVDFEPKPLVEEPQNLIKONCELFQELGEYKFNALLVRYTKK	414
Db	360	ELLEKCFQTNPLEFCQDKEEELQKYIQESQALAKRSGCLFQKLGEYYLQNEFLVAYTKK	419
Qy	415	VPQVSTFTLVEVSRNLKVGSKCKHPEAKMPCAEEDYLSVVLNQLCVLHEKTPVSDRVT	474
Db	420	APQLTSSSELMAITRKMAATAATCCQLSEDKILACGEGAADIIIGHLCIRHEMTFVNPVG	479
Qy	475	KCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTAALVELVK	534
Db	480	QCTSSVANRRPCFSSLVVDETYVPPAFSDDKFIHFKDLCOAQGVALQRMKQBEFLINLVK	539
Qy	535	HKPKATKEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAAASQAALGL	585
Db	540	KPQITEQLEALIAFSGLEKCCQGEQVCFABEGQKLISKTGAAALGV	590

Search completed: October 17, 2005, 08:29:42  
Job time : 47 secs

**THIS PAGE BLANK (08/17/10)**

Fused

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2005, 08:25:57 ; Search time 169 Seconds  
(without alignments)  
1770.915 Million cell updates/sec

Title: US-10-933-523-18

Perfect score: 3785

Sequence: 1 DAKHSEVAHRFKDLGLENFX.....leqlrelrigsgvlgcvhs 718

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.1

Searched: 1860064 seqs, 416830855 residues

Total number of hits satisfying chosen parameters: 1860064

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 100 summaries

Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pdb.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pdb.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pdb.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pdb.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pdb.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pdb.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pdb.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pdb.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pdb.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pdb.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pdb.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pdb.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pdb.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pdb.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pdb.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pdb.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pdb.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US11A\_PUBCOMB.pdb.\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pdb.\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pdb.\*
- 21: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pdb.\*
- 22: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pdb.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3156.5	83.4	763	15	US-10-609-346-2
2	3153.4	83.3	838	17	US-10-775-180-590
3	3153.4	83.3	838	18	US-10-775-204-1591
4	3152.8	83.3	876	17	US-10-775-180-82
5	3152.8	83.3	876	18	US-10-775-204-236
6	3152.3	83.3	819	17	US-10-775-180-603
7	3152.3	83.3	819	17	US-10-775-180-605
8	3152.3	83.3	819	18	US-10-775-204-1614
9	3152.3	83.3	819	18	US-10-775-204-1616
10	3152.2	83.3	794	18	US-10-775-204-1657
11	3152.2	83.3	809	18	US-10-775-204-1548
12	3152.2	83.3	809	18	US-10-775-204-1548
13	3152.2	83.3	809	18	US-10-775-204-1548
14	3152.2	83.3	809	18	US-10-775-204-1548
15	3152.2	83.3	809	18	US-10-775-204-1548
16	3152.2	83.3	809	18	US-10-775-204-1548
17	3152.2	83.3	809	18	US-10-775-204-1548
18	3152.2	83.3	809	18	US-10-775-204-1548
19	3152.2	83.3	809	18	US-10-775-204-1548
20	3152.2	83.3	809	18	US-10-775-204-1548
21	3152.2	83.3	809	18	US-10-775-204-1548
22	3152.2	83.3	809	18	US-10-775-204-1548
23	3152.2	83.3	809	18	US-10-775-204-1548
24	3152.2	83.3	809	18	US-10-775-204-1548
25	3152.2	83.3	809	18	US-10-775-204-1548
26	3152.2	83.3	809	18	US-10-775-204-1548
27	3152.2	83.3	809	18	US-10-775-204-1548
28	3152.2	83.3	809	18	US-10-775-204-1548
29	3152.2	83.3	809	18	US-10-775-204-1548
30	3152.2	83.3	809	18	US-10-775-204-1548
31	3152.2	83.3	809	18	US-10-775-204-1548
32	3152.2	83.3	809	18	US-10-775-204-1548
33	3152.2	83.3	809	18	US-10-775-204-1548
34	3152.2	83.3	809	18	US-10-775-204-1548
35	3152.2	83.3	809	18	US-10-775-204-1548
36	3152.2	83.3	809	18	US-10-775-204-1548
37	3149.6	83.2	768	18	US-10-775-204-1521
38	3149.6	83.2	768	18	US-10-775-204-1521
39	3149.6	83.2	768	18	US-10-775-204-1523
40	3149.6	83.2	768	18	US-10-775-204-1660
41	3149.6	83.2	768	18	US-10-775-204-1661
42	3149.6	83.2	768	18	US-10-775-204-1662
43	3149.6	83.2	768	18	US-10-775-204-401
44	3149.2	83.2	892	18	US-10-775-204-1658
45	3146.5	83.1	820	18	US-10-775-204-228
46	3145.8	83.1	894	18	US-10-775-204-261
47	3145.7	83.1	864	17	US-10-775-180-73
48	3145.7	83.1	864	17	US-10-775-204-220
49	3145.1	83.1	747	17	US-10-775-180-602
50	3145.1	83.1	747	17	US-10-775-180-604
51	3145.1	83.1	747	18	US-10-775-204-1613
52	3145.1	83.1	747	18	US-10-775-204-1615
53	3144.9	83.1	870	17	US-10-775-180-81
54	3144.9	83.1	870	18	US-10-775-204-235
55	3144.9	83.1	879	17	US-10-775-180-77
56	3144.9	83.1	879	18	US-10-775-204-224
57	3144.4	83.1	864	18	US-10-775-204-296
58	3142.1	83.0	778	18	US-10-775-204-1546
59	3141.9	83.0	786	18	US-10-775-204-1534
60	3140.7	83.0	724	17	US-10-928-305-5
61	3140.7	83.0	742	18	US-10-775-204-243
62	3140.7	83.0	742	18	US-10-775-204-259
63	3140.7	83.0	742	18	US-10-775-204-320
64	3140.6	83.0	778	18	US-10-775-204-1544
65	3140.6	83.0	779	18	US-10-775-204-1545
66	3140.6	83.0	780	18	US-10-775-204-1194
67	3140.6	83.0	783	18	US-10-775-204-226
68	3140.6	83.0	783	18	US-10-775-204-1585
69	3140.6	83.0	951	18	US-10-775-204-404
70	3140.6	83.0	951	18	US-10-775-204-409
71	3138.7	82.9	771	18	US-10-775-204-1582
72	3138.7	82.9	771	18	US-10-775-204-1581
73	3138.5	82.9	746	15	US-10-609-346-6
74	3138.5	82.9	774	17	US-10-775-180-124
75	3138.5	82.9	774	17	US-10-775-180-452
76	3138.5	82.9	774	18	US-10-775-204-402
77	3138.5	82.9	774	18	US-10-775-204-1300
78	3138.4	82.9	1015	17	US-10-775-180-591
79	3138.4	82.9	1015	17	US-10-775-204-1532
80	3137.9	82.9	671	17	US-10-775-180-625
81	3137.9	82.9	671	17	US-10-775-204-1645
82	3137.7	82.9	774	17	US-10-775-180-455
83	3137.7	82.9	774	17	US-10-775-204-1309
84	3137.6	82.9	759	15	US-10-609-346-8

85 3137 82.9 770 18 US-10-775-204-1583 Sequence 1583, Ap  
86 3137 82.9 773 18 US-10-775-204-1584 Sequence 1584, Ap  
87 3134.1 82.8 774 17 US-10-775-180-470 Sequence 470, App  
88 3134.1 82.8 774 18 US-10-775-204-1324 Sequence 1324, Ap  
89 3133.7 82.8 774 17 US-10-775-180-475 Sequence 475, App  
90 3133.2 82.8 774 18 US-10-775-204-1329 Sequence 1329, Ap  
91 3133.2 82.8 1010 17 US-10-775-180-556 Sequence 556, App  
92 3133.2 82.8 1010 17 US-10-775-180-557 Sequence 557, App  
93 3133.2 82.8 1010 17 US-10-775-180-558 Sequence 558, App  
94 3133.2 82.8 1010 17 US-10-775-180-559 Sequence 559, App  
95 3133.2 82.8 1010 17 US-10-775-180-560 Sequence 560, App  
96 3133.2 82.8 1010 17 US-10-775-180-561 Sequence 561, App  
97 3133.2 82.8 1010 18 US-10-775-204-1527 Sequence 1527, Ap  
98 3133.2 82.8 1010 18 US-10-775-204-1528 Sequence 1528, Ap  
99 3133.2 82.8 1010 18 US-10-775-204-1529 Sequence 1529, Ap  
100 3133.2 82.8 1010 18 US-10-775-204-1531 Sequence 1531, Ap

ALIGNMENTS

RESULT 1  
US-10-609-346-2  
; Sequence 2, Application US/10609346  
; Publication No. US20040063635A1  
; GENERAL INFORMATION:  
; APPLICANT: Yu, Zailin  
; APPLICANT: Fu, Yan  
; TITLE OF INVENTION: RECOMBINANT HUMAN ALBUMIN FUSION PROTEINS WITH LONG-LASTING BIOLOGICAL ACTIVITY  
; FILE REFERENCE: ZYU-0603  
; CURRENT APPLICATION NUMBER: US/10/609,346  
; CURRENT FILING DATE: 2003-06-26  
; PRIOR APPLICATION NUMBER: US 60/392,948  
; PRIOR FILING DATE: 2002-07-01  
; NUMBER OF SEQ ID NOS: 40  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 763  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: HSA-hil-11  
US-10-609-346-2

Query Match 83.4%; Score 3156.5; DB 15; Length 763;  
Best Local Similarity 80.9%; Pred. No. 1.5e-75;  
Matches 619; Conservative 11; Mismatches 40; Indels 95; Gaps 9;  
QY 1 DAHKSVAHRFKDLGLENFKALVLIAPAOYLOQCPEDHVKLVNEVTEFAKTCVADESAB 60  
DB 1 DAHKSVAHRFKDLGLENFKALVLIAPAOYLOQCPEDHVKLVNEVTEFAKTCVADESAB 60  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFQLQKDDNPNLRLVRPEV 120  
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFQLQKDDNPNLRLVRPEV 120  
QY 121 DVNCTAFHNEETFLKYLVEIARRHPYFVAPLLFFAKRYKAATFECCQAADKAACLLP 180  
DB 121 DVNCTAFHNEETFLKYLVEIARRHPYFVAPLLFFAKRYKAATFECCQAADKAACLLP 180  
QY 181 KLDLDEGKASSAKORLKASLOKGERAFKAWAVARLSORPPKAEFAEVSKLVTDLTK 240  
DB 181 KLDLDEGKASSAKORLKASLOKGERAFKAWAVARLSORPPKAEFAEVSKLVTDLTK 240  
QY 241 VHTCCGDLLECCADRADLAKYICENQDSISSKLKCECKEPLEKSHCIAEVNDEMPPA 300  
DB 241 VHTCCGDLLECCADRADLAKYICENQDSISSKLKCECKEPLEKSHCIAEVNDEMPPA 300  
QY 301 DLPSLAADFVSKOVCKNAEAKDVFGLMFLYEARRHPDYSVLLLRLLRAKTYETTLK 360  
DB 301 DLPSLAADFVSKOVCKNAEAKDVFGLMFLYEARRHPDYSVLLLRLLRAKTYETTLK 360

QY 361 CAAADPHECYAKVDFDEKPLVEEPQNLIKQNCLEFQOLGEYKFNQALLVRYTKVPQVST 420  
DB 361 CAAADPHECYAKVDFDEKPLVEEPQNLIKQNCLEFQOLGEYKFNQALLVRYTKVPQVST 420  
QY 421 PTLVEVSRNLGKVGSKCKCKHPKAMPKPCAEEDYLSVVNLQICVLHEKTPVSDRVTKCCTES 480  
DB 421 PTLVEVSRNLGKVGSKCKCKHPKAMPKPCAEEDYLSVVNLQICVLHEKTPVSDRVTKCCTES 480  
QY 481 LVNRRPCFSALEVDEITYVPKFNATFTPHADICTLSEKEROIKKOTALVELVKKHKPKAT 540  
DB 481 LVNRRPCFSALEVDEITYVPKFNATFTPHADICTLSEKEROIKKOTALVELVKKHKPKAT 540  
QY 541 KEQLKAVMDDFAAFEVKCKKADDDKTCFAEKGKLVAAASQAALGLMSPLRLEVPCHALPQ 600  
DB 541 KEQLKAVMDDFAAFEVKCKKADDDKTCFAEKGKLVAAASQAALGLPQP-----PP 590  
QY 601 G---LSP-----GOVIVRGLVLOEPKHFTVSRDQ----- 628  
DB 591 GPPRVSPDRAELDSTVLLTRSL-LADTRQLAAQLRDKFPADGDHNLDSLPTLAMSAGAL 649  
QY 629 -AAHAPVTLRASPAD-----RTLQWISRWGKKLIS----- 658  
DB 650 GALQUPGVLTURADLLSYLRHVQMLRRAGGSSSLKTLPELGTLOARLDRLLRLQLLMS 709  
QY 659 -----APFLFYQP-----RPFVLLLFQEGGLKALN 685  
DB 710 RLALPQPPDPAPPPLAPSSAWGGIRAAHAIL-----GGLHLTLD 750

RESULT 2

US-10-775-180-590  
; Sequence 590, Application US/10775180  
; Publication No. US20050054570A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, Craig A.  
; APPLICANT: Haseltine, William A.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF574  
; CURRENT APPLICATION NUMBER: US/10/775,180  
; CURRENT FILING DATE: 2004-02-11  
; PRIOR APPLICATION NUMBER: PCT/US02/40892  
; PRIOR FILING DATE: 2002-12-23  
; PRIOR APPLICATION NUMBER: 60/341,811  
; PRIOR FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/360,000  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: 60/378,950  
; PRIOR FILING DATE: 2002-03-10  
; PRIOR APPLICATION NUMBER: 60/398,008  
; PRIOR FILING DATE: 2002-07-24  
; PRIOR APPLICATION NUMBER: 60/411,355  
; PRIOR FILING DATE: 2002-09-18  
; PRIOR APPLICATION NUMBER: 60/414,984  
; PRIOR FILING DATE: 2002-10-02  
; PRIOR APPLICATION NUMBER: 60/417,611  
; PRIOR FILING DATE: 2002-10-11  
; PRIOR APPLICATION NUMBER: 60/420,246  
; PRIOR FILING DATE: 2002-10-23  
; PRIOR APPLICATION NUMBER: 60/423,623  
; PRIOR FILING DATE: 2002-11-05  
; Remaining prior application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 858  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 590  
; LENGTH: 838  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-775-180-590

Query Match 83.3%; Score 3153.4; DB 17; Length 838;  
Best Local Similarity 75.0%; Pred. No. 2.2e-75;  
Matches 624; Conservative 15; Mismatches 37; Indels 156; Gaps 12;

```
QY 1 DAHSEVAHRFKDLGSENFKALVLIAPAFYQLOQCPPEHNVKLVNVEVTEPAKTCVADESAE 60
Db 25 DAHSEVAHRFKDLGSENFKALVLIAPAFYQLOQCPPEHNVKLVNVEVTEPAKTCVADESAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144
QY 121 DVNCTAFHNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 180
Db 145 DVNCTAFHNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 204
QY 181 KLDELDEGKASSAKQRLKASLOKGERAFKAWAVARLSQRPKAPFAEVSKLVTDLTK 240
Db 205 KLDELDEGKASSAKQRLKASLOKGERAFKAWAVARLSQRPKAPFAEVSKLVTDLTK 264
QY 241 VHTTECHGDLLECCADDRADLAKYICENODSISSEKKECKEPLLEKSHCIAEVNDEMPA 300
Db 265 VHTTECHGDLLECCADDRADLAKYICENODSISSEKKECKEPLLEKSHCIAEVNDEMPA 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYFYARRHPDYSVVLLRLAKTYETTLK 360
Db 325 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYFYARRHPDYSVVLLRLAKTYETTLK 384
QY 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKVQVST 420
Db 385 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKVQVST 444
QY 421 PTLVEVSRNLGKVGSKCKCHPEAKRMPCAEDYLSVNLNQLCVLHEKTPVSDRVTKCCTES 480
Db 445 PTLVEVSRNLGKVGSKCKCHPEAKRMPCAEDYLSVNLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDVETVPKFEFNAETFTFHADICTLSEKERQIKKOTALVELVKHKPKAT 540
Db 505 LVNRRPCFSALEVDVETVPKFEFNAETFTFHADICTLSEKERQIKKOTALVELVKHKPKAT 564
QY 541 KEQLKAVMDDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGLM----- 586
Db 565 KEQLKAVMDDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGLMDLLQLFLAPLVLL 624
QY 587 -----SPRLEV----- 592
Db 625 SGMGATGTLRTSLDPSLEIYKKMFVKRRREQLLAKNLALQINDIHOQYKILDVMLKGLFK 684
QY 593 -----PCSHALPGLSPGOVIVRGVLQEPKHFTVSLRDOAAHAPVTLRASFADR 643
Db 685 VLDSRTVLTAAADVLPQPP-----QDEK-----LKDAFHV-VENTAFGDV 727
QY 644 TLQ-----W--ISRWGQKGLISAPFLFYQPFVPEVLLLFQEGGLK--LALN 685
Db 728 VLAPPRIVHYFDHNSNWNLLIRWG-----IS-----FCNQTGVNQGPHSPILSLM 774
QY 686 GQGLG-----ATSNQOAL-----EQRLER 706
Db 775 AQELGISSEKDSNFQNPFPKIDRTEFIPSTDPFQKALREERKREKREIR 826
RESULT 3
US-10-775-204-1591
; Sequence 1591, Application US/10775204
; Publication No. US20050186664A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseeltine, William A.
; APPLICANT: Balance, David J.
; APPLICANT: Turner, Andrew J.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF564
; CURRENT APPLICATION NUMBER: US/10/775,204
; CURRENT FILING DATE: 2004-02-11
; PRIOR APPLICATION NUMBER: 60/341,811
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/360,000
```

```
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/378,950
; PRIOR FILING DATE: 2002-03-10
; PRIOR APPLICATION NUMBER: 60/398,008
; PRIOR FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: 60/411,355
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/414,984
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: 60/417,611
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: 60/420,246
; PRIOR FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: 60/423,623
; PRIOR FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: 60/351,360
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2222
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1591
; LENGTH: 838
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-775-204-1591
Query Match 83.3%; Score 3153.4; DB 18; Length 838;
Best Local Similarity 75.0%; Pred. No. 2.2e-75;
Matches 624; Conservative 15; Mismatches 37; Indels 156; Gaps 12;
QY 1 DAHSEVAHRFKDLGSENFKALVLIAPAFYQLOQCPPEHNVKLVNVEVTEPAKTCVADESAE 60
Db 25 DAHSEVAHRFKDLGSENFKALVLIAPAFYQLOQCPPEHNVKLVNVEVTEPAKTCVADESAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144
QY 121 DVNCTAFHNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 180
Db 145 DVNCTAFHNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 204
QY 181 KLDELDEGKASSAKQRLKASLOKGERAFKAWAVARLSQRPKAPFAEVSKLVTDLTK 240
Db 205 KLDELDEGKASSAKQRLKASLOKGERAFKAWAVARLSQRPKAPFAEVSKLVTDLTK 264
QY 241 VHTTECHGDLLECCADDRADLAKYICENODSISSEKKECKEPLLEKSHCIAEVNDEMPA 300
Db 265 VHTTECHGDLLECCADDRADLAKYICENODSISSEKKECKEPLLEKSHCIAEVNDEMPA 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYFYARRHPDYSVVLLRLAKTYETTLK 360
Db 325 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYFYARRHPDYSVVLLRLAKTYETTLK 384
QY 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKVQVST 420
Db 385 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKVQVST 444
QY 421 PTLVEVSRNLGKVGSKCKCHPEAKRMPCAEDYLSVNLNQLCVLHEKTPVSDRVTKCCTES 480
Db 445 PTLVEVSRNLGKVGSKCKCHPEAKRMPCAEDYLSVNLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDVETVPKFEFNAETFTFHADICTLSEKERQIKKOTALVELVKHKPKAT 540
Db 505 LVNRRPCFSALEVDVETVPKFEFNAETFTFHADICTLSEKERQIKKOTALVELVKHKPKAT 564
QY 541 KEQLKAVMDDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGLM----- 586
Db 565 KEQLKAVMDDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGLMDLLQLFLAPLVLL 624
QY 587 -----SPRLEV----- 592
Db 625 SGMGATGTLRTSLDPSLEIYKKMFVKRRREQLLAKNLALQINDIHOQYKILDVMLKGLFK 684
```

QY 593 -----PCSHALPOGLSPGVIIIRGLVQLQPKHFTVSLRDOAAHAPVTLRASFADR 643  
DB 685 VLEDSRTVLTADVLDPGFP-----QDEK-----LKDAFSHV-VENTAFPGDV 727  
QY 644 TLQ-----W-----ISRWGQKKUISAPFLFYQRFVFLVLLFOEGGLK--LAIN 685  
DB 728 VLRFPRIHVHYEDHNSWNLLIRWG-----IS-----FCNQTGVFNQGHSPILSLM 774  
QY 686 GQGLG-----ATSMNQOAL-----BOLRELR 706  
DB 775 AQELGISEKSNFQNPFKIDRTFIPSTDPFQKALREERKRKKEIR 826

RESULT 4

US-10-775-180-82  
; Sequence 82, Application US/10775180  
; Publication No. US20050054570A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, Craig A.  
; APPLICANT: Haseltine, William A.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF574  
; CURRENT APPLICATION NUMBER: US/10/775,180  
; CURRENT FILING DATE: 2004-02-11  
; PRIOR APPLICATION NUMBER: PCT/US02/40892  
; PRIOR FILING DATE: 2002-12-23  
; PRIOR APPLICATION NUMBER: 60/341,811  
; PRIOR FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/360,000  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: 60/378,950  
; PRIOR FILING DATE: 2002-05-10  
; PRIOR APPLICATION NUMBER: 60/398,008  
; PRIOR FILING DATE: 2002-07-24  
; PRIOR APPLICATION NUMBER: 60/411,355  
; PRIOR FILING DATE: 2002-09-18  
; PRIOR APPLICATION NUMBER: 60/414,984  
; PRIOR FILING DATE: 2002-10-02  
; PRIOR APPLICATION NUMBER: 60/417,611  
; PRIOR FILING DATE: 2002-10-11  
; PRIOR APPLICATION NUMBER: 60/420,246  
; PRIOR FILING DATE: 2002-10-23  
; PRIOR APPLICATION NUMBER: 60/423,623  
; PRIOR FILING DATE: 2002-11-05  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 858  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 82  
; LENGTH: 876  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-775-180-82

Query Match 83.3%; Score 3152.8; DB 17; Length 876;  
Best Local Similarity 70.4%; Pred. No. 2.5e-75;  
Matches 622; Conservative 13; Mismatches 36; Indels 212; Gaps 10;

QY 1 DAHSEVAHRFKDLGENFKALVLIAPQYLOQCPEDHVKLVNVEFPAKTCVADESAE 60  
DB 25 DAHSEVAHRFKDLGENFKALVLIAPQYLOQCPEDHVKLVNVEFPAKTCVADESAE 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 144  
QY 121 DWCTAFDNEETFLKKLYEIAARRHPYFYPAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180  
DB 145 DWCTAFDNEETFLKKLYEIAARRHPYFYPAPPELLFFAKRYKAAFTCCQAAADKAACLLP 204  
QY 181 KLDELDEGKASSAKORLKCSAQKFGGERAFKAWAVARLSQRPPKAEFVSKLVTDLTK 240  
DB 205 KLDELDEGKASSAKORLKCSAQKFGGERAFKAWAVARLSQRPPKAEFVSKLVTDLTK 264

QY 241 VHTTECHGDLLESCADRADLAKYICENODSISSKLKECCEKPLLEKSHCIAEVENDEMPA 300  
DB 265 VHTTECHGDLLESCADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 324  
QY 301 DLPSLAADFVESKDVCKNVAEAKDVLGMLFYEARHHPDYSVVLLLRLLAKTYETTTLEKC 360  
DB 325 DLPSLAADFVESKDVCKNVAEAKDVLGMLFYEARHHPDYSVVLLLRLLAKTYETTTLEKC 384  
QY 361 CAAADPHCYAKVDFEPKPLVEEPQNLKQNCLEFQOLGEYFQNALVRYTKKVPQVST 420  
DB 385 CAAADPHCYAKVDFEPKPLVEEPQNLKQNCLEFQOLGEYFQNALVRYTKKVPQVST 444  
QY 421 PTLVEVSRNLGKVGSKCKHPEAKMPCAEADYLSVVLNQLCVLHETKPVSDRVTKCCTES 480  
DB 445 PTLVEVSRNLGKVGSKCKHPEAKMPCAEADYLSVVLNQLCVLHETKPVSDRVTKCCTES 504  
QY 481 LVNRRPCFSALEVDYTYVPKEFNAETFTFHADICTLSEKERQIKKQATALVELVGHKPKAT 540  
DB 505 LVNRRPCFSALEVDYTYVPKEFNAETFTFHADICTLSEKERQIKKQATALVELVGHKPKAT 564  
QY 541 KEQLXAVMDDFAAVFEKCKCKADDDKTCFAEBEGKLVAAASQAALGLMSRLEV----- 592  
DB 565 KEQLXAVMDDFAAVFEKCKCKADDDKTCFAEBEGKLVAAASQAALGL-----LEVAETPTYPW 620  
QY 593 -----PCSHALPOGLSPGVIIIRGLVQLQPKHFT----- 622  
DB 621 RDAETGERLVCAQCPGTFVQRPCTRDSTTCGP-----CPRHYTQFWNYLER 669  
QY 623 -----VSLRDQAAHA----- 632  
DB 670 CRYCNVLGGEREEARACHATHNRACRGTGFFAHAGFCLEHASCPPGAGVIAPGTSPQN 729  
QY 633 -----PVTLRAS----- 639  
DB 730 TOCQPCPPGTFSSASSSSBQCQPHRNCNTALGLALNVPGSSSHDTLCTSGTGPPLSTRVPG 789  
QY 640 -----FADRTLQWISR-----WG-----QKGLISAPFLFYPO 666  
DB 790 AECEERAVIDFVAFODISIKRLQRLLOALEAPEGHGPTPRAGRAALQLKL-----R 840  
QY 667 RFEVLLLFQEGGLKALNGQGLGATSMNQQAELQELRELIRSG 709  
DB 841 RRLTELLGAQDGALLVRL-----LQALVARMPG 869

RESULT 5

US-10-775-204-236  
; Sequence 236, Application US/10775204  
; Publication No. US20050186664A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, Craig A.  
; APPLICANT: Haseltine, William A.  
; APPLICANT: Balance, David J.  
; APPLICANT: Turner, Andrew J.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF564  
; CURRENT APPLICATION NUMBER: US/10/775,204  
; CURRENT FILING DATE: 2004-02-11  
; PRIOR APPLICATION NUMBER: 60/341,811  
; PRIOR FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/360,000  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: 60/378,950  
; PRIOR FILING DATE: 2002-05-10  
; PRIOR APPLICATION NUMBER: 60/398,008  
; PRIOR FILING DATE: 2002-07-24  
; PRIOR APPLICATION NUMBER: 60/411,355  
; PRIOR FILING DATE: 2002-09-18  
; PRIOR APPLICATION NUMBER: 60/414,984  
; PRIOR FILING DATE: 2002-10-02  
; PRIOR APPLICATION NUMBER: 60/417,611  
; PRIOR FILING DATE: 2002-10-11

1 PRIOR APPLICATION NUMBER: 60/420,246  
2 PRIOR FILING DATE: 2002-10-23  
3 PRIOR APPLICATION NUMBER: 60/423,623  
4 PRIOR FILING DATE: 2002-11-05  
5 PRIOR APPLICATION NUMBER: 60/351,360  
6 PRIOR FILING DATE: 2002-01-28  
7 Remaining Prior Application data removed - See File Wrapper or PALM.  
8 NUMBER OF SEQ ID NOS: 2222  
9 SOFTWARE: Patent In Ver. 2.0  
10 SEQ ID NO 236  
11 LENGTH: 876  
12 TYPE: PRT  
13 ORGANISM: Homo sapiens  
14 US-10-775-204-236

Query Match 83.3%; Score 3152.8; DB 18; Length 876;  
Best Local Similarity 70.4%; Pred. No. 2.5e-75;  
Matches 622; Conservative 13; Mismatches 36; Indels 212; Gaps 10;

QY 1 DAHSEVAHRFKDGBENFKALVLIAPAQYLOQCPEDHVKLVNEVTEFAKTCVADESAAE 60  
DB 25 DAHSEVAHRFKDGBENFKALVLIAPAQYLOQCPEDHVKLVNEVTEFAKTCVADESAAE 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144  
QY 121 DVNCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFAKRYKAFTCCCAADKAACLLP 180  
DB 145 DVNCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFAKRYKAFTCCCAADKAACLLP 204  
QY 181 KLDELDEGKASSAKORLKCASLQKFGERAFKAWAVARLSQRPFPKAEFAEVSCLVTDLT 240  
DB 205 KLDELDEGKASSAKORLKCASLQKFGERAFKAWAVARLSQRPFPKAEFAEVSCLVTDLT 264  
QY 241 VHTCCGHDLLCADDRAADIAKYICENQDSISSKLKECCCKPILKSHCHIAEVENDEMPA 300  
DB 265 VHTCCGHDLLCADDRAADIAKYICENQDSISSKLKECCCKPILKSHCHIAEVENDEMPA 324  
QY 301 DLPSLAADFVESKDVCKNVAEKDVLGMFLYVARRHPDYSVVLRLAKTVEYTTLEK 360  
DB 325 DLPSLAADFVESKDVCKNVAEKDVLGMFLYVARRHPDYSVVLRLAKTVEYTTLEK 384  
QY 361 CAADPHCEYAKVDFEKLVEEPQNLIKONCELFQGLGYKFNALLVRYTKVPQVST 420  
DB 385 CAADPHCEYAKVDFEKLVEEPQNLIKONCELFQGLGYKFNALLVRYTKVPQVST 444  
QY 421 PTLVEVSRNLGKSGCKCKHPEAKRMPCABEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 445 PTLVEVSRNLGKSGCKCKHPEAKRMPCABEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 504  
QY 481 LVNRRPCFSALEVDVTVPKFNAETTFHADICTLSEKROIKKOTALVELVKGHPKAT 540  
DB 505 LVNRRPCFSALEVDVTVPKFNAETTFHADICTLSEKROIKKOTALVELVKGHPKAT 564  
QY 541 KEQLKAWMDFFAAVFEKCCCKADKTCFAEBGKGLVAASQAALGMSPLREV----- 592  
DB 565 KEQLKAWMDFFAAVFEKCCCKADKTCFAEBGKGLVAASQAALG-----LEVAETPTYPW 620  
QY 593 -----PCSHALPOGLSPGQVIVRGLVLQBPQKHT----- 622  
DB 621 RDAETGERLVCAQCPGTFVQRCRDRSPPTTCG-----CPRHYTQWNYLER 669  
QY 623 -----VSLRDQAHA----- 632  
DB 670 CRYCNVLGGEREEARACHATHNACRCRTGFFAHAGFCLEHASCPPGAGVIAPGTPSQN 729  
QY 633 -----PVTLRAS----- 639  
DB 730 TQOCPPGPTFSASSSSSEQPHNCTALGLALNVPGSSSHDILTCTGCTGFPPLSTRVPG 789  
QY 640 -----FADRTLOWISR-----WG-----QKCLISAPFLPYPP 666

DB 790 ASECERAVIDFVAFQDISIKRLQRLLOALEAPEGWGTTPRAGRAALQLKL-----R 840  
QY 667 RFPEVILLFQEGGLKALNGOGLGATSMNQQAELQRLRLRISG 709  
DB 841 RRLTELLGAGDQALLVRL-----LQALRVARMPG 869

RESULT 6  
US-10-775-180-603  
; Sequence 603, Application US/10775180  
; Publication No. US20050054570A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, Craig A.  
; APPLICANT: Haseltine, William A.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF574  
; CURRENT APPLICATION NUMBER: US/10/775,180  
; CURRENT FILING DATE: 2004-02-11  
; PRIOR APPLICATION NUMBER: PCT/US02/40892  
; PRIOR FILING DATE: 2002-12-23  
; PRIOR APPLICATION NUMBER: 60/341,811  
; PRIOR FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/360,000  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: 60/378,950  
; PRIOR FILING DATE: 2002-05-10  
; PRIOR APPLICATION NUMBER: 60/398,008  
; PRIOR FILING DATE: 2002-07-24  
; PRIOR APPLICATION NUMBER: 60/411,355  
; PRIOR FILING DATE: 2002-09-18  
; PRIOR APPLICATION NUMBER: 60/414,984  
; PRIOR FILING DATE: 2002-10-02  
; PRIOR APPLICATION NUMBER: 60/417,611  
; PRIOR FILING DATE: 2002-10-11  
; PRIOR APPLICATION NUMBER: 60/420,246  
; PRIOR FILING DATE: 2002-10-23  
; PRIOR APPLICATION NUMBER: 60/423,623  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 858  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 603  
; LENGTH: 819  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-775-180-603

Query Match 83.3%; Score 3152.3; DB 17; Length 819;  
Best Local Similarity 76.8%; Pred. No. 2.3e-75;  
Matches 623; Conservative 16; Mismatches 37; Indels 137; Gaps 12;

QY 1 DAHSEVAHRFKDGBENFKALVLIAPAQYLOQCPEDHVKLVNEVTEFAKTCVADESAAE 60  
DB 25 DAHSEVAHRFKDGBENFKALVLIAPAQYLOQCPEDHVKLVNEVTEFAKTCVADESAAE 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144  
QY 121 DVNCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFAKRYKAFTCCCAADKAACLLP 180  
DB 145 DVNCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFAKRYKAFTCCCAADKAACLLP 204  
QY 181 KLDELDEGKASSAKORLKCASLQKFGERAFKAWAVARLSQRPFPKAEFAEVSCLVTDLT 240  
DB 205 KLDELDEGKASSAKORLKCASLQKFGERAFKAWAVARLSQRPFPKAEFAEVSCLVTDLT 264  
QY 241 VHTCCGHDLLCADDRAADIAKYICENQDSISSKLKECCCKPILKSHCHIAEVENDEMPA 300  
DB 265 VHTCCGHDLLCADDRAADIAKYICENQDSISSKLKECCCKPILKSHCHIAEVENDEMPA 324  
QY 301 DLPSLAADFVESKDVCKNVAEKDVLGMFLYVARRHPDYSVVLRLAKTVEYTTLEK 360



Db 325 DLPSLAADFVSKGVKQVNAEAKDVFGLMFLEYARRHPDYSVVLRLIAKTYETTLEKC 384  
QY 361 CAAADPHECYAKVDFDFKPLVREPQNLIKONCELPQOLGEYKFNALLVRYTKVPQVST 420  
Db 385 CAAADPHECYAKVDFDFKPLVREPQNLIKONCELPQOLGEYKFNALLVRYTKVPQVST 444  
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHKTVPVSDRVTKCCTES 480  
Db 445 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHKTVPVSDRVTKCCTES 504  
QY 481 LVNRRPCFSALEVDVETVYKFNATFTFHADICTLSEKEROIKKOTALVELVGHKPKAT 540  
Db 505 LVNRRPCFSALEVDVETVYKFNATFTFHADICTLSEKEROIKKOTALVELVGHKPKAT 564  
QY 541 KEQLKAVMDDDFAAFVEKCKCADDKETCFABEGKKLVAASQAALGL- - - - -MSPRLEV 592  
Db 565 KEQLKAVMDDDFAAFVEKCKCADDKETCFABEGKKLVAASQAALGLATGLTSLDPSLEI 624  
QY 593 - - - - -ATSMNQAL- - - - -EQLREL 706  
Db 625 YKMFVKKRRQQLLAKNLQALNDIHQQYKILDMVLMKGLFKVLEDSRTVLTAAADVLPDGP 684  
QY 603 SPGQVLIIVRGLVLOPKHFTVSLRDOAAHAPVTLRASPADRTLQ- - - - -W- 647  
Db 685 FP- - - - -QDEK- - - - -LKDAFSHV-VENTAFGDDVVLFRPRIVHYVFDHNSWN 727  
QY 648 -ISRMGQKKLISAPFLFYPPQRFVLELLFQEGGLK--LALNGQGLG- - - - - 690  
Db 728 LLIRWG- - - - -IS- - - - -FCNQTGVFNQGHSPHLSLMAQELGISEKDSNFQNPFKI 774  
QY 691 - - - - -ATSMNQAL- - - - -EQLREL 706  
Db 775 DRTEFIPSTDPPQKALREEEKRKRKEIR 807

RESULT 7

US-10-775-180-605  
; Sequence 605, Application US/10775180  
; Publication No. US20050054570A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, Craig A.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF574  
; CURRENT APPLICATION NUMBER: US/10/775,180  
; CURRENT FILING DATE: 2004-02-11  
; PRIOR APPLICATION NUMBER: PCT/US02/40892  
; PRIOR FILING DATE: 2002-12-23  
; PRIOR APPLICATION NUMBER: 60/341,811  
; PRIOR FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/360,000  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: 60/378,950  
; PRIOR FILING DATE: 2002-05-10  
; PRIOR APPLICATION NUMBER: 60/398,008  
; PRIOR FILING DATE: 2002-07-24  
; PRIOR APPLICATION NUMBER: 60/411,355  
; PRIOR FILING DATE: 2002-09-18  
; PRIOR APPLICATION NUMBER: 60/414,984  
; PRIOR FILING DATE: 2002-10-02  
; PRIOR APPLICATION NUMBER: 60/417,611  
; PRIOR FILING DATE: 2002-10-11  
; PRIOR APPLICATION NUMBER: 60/420,246  
; PRIOR FILING DATE: 2002-10-23  
; PRIOR APPLICATION NUMBER: 60/423,623  
; PRIOR FILING DATE: 2002-11-05  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 858  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 605  
; LENGTH: 819  
; TYPE: PRT  
; ORGANISM: Homo sapiens

US-10-775-180-605  
Query Match 83.3%; Score 3152.3; DB 17; Length 819;  
Best Local Similarity 76.6%; Pred. No. 2.3e-75;  
Matches 623; Conservative 16; Mismatches 37; Indels 137; Gaps 12;  
QY 1 DAHKSEVAHRFKDLGEENFKALVLIATFAQYLQOCCPFEDHVKLVNVEVTEFAKTCVADESAB 60  
Db 25 DAHKSEVAHRFKDLGEENFKALVLIATFAQYLQOCCPFEDHVKLVNVEVTEFAKTCVADESAB 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECPLQHKDDNPNLPRLVREV 120  
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECPLQHKDDNPNLPRLVREV 144  
QY 121 DVMCATFADHNBETFLKKYLYETARRHPYFYAPPELLFFAKRYKAAFTTECCOAAADKAAACLLP 180  
Db 145 DVMCATFADHNBETFLKKYLYETARRHPYFYAPPELLFFAKRYKAAFTTECCOAAADKAAACLLP 204  
QY 181 KLDELDRBEGKASSAKQRLKCAASLQKFGRAFKAWAVARLSQRFPPKAEFAEVSCLVTDLTK 240  
Db 205 KLDELDRBEGKASSAKQRLKCAASLQKFGRAFKAWAVARLSQRFPPKAEFAEVSCLVTDLTK 264  
QY 241 VHTECHGDLLEACADDRADLAKYICENQDSISSKKECEKPLLEKSHCIAEVENDEMPA 300  
Db 265 VHTECHGDLLEACADDRADLAKYICENQDSISSKKECEKPLLEKSHCIAEVENDEMPA 324  
QY 301 DLPSLAADFVSKGVKQVNAEAKDVFGLMFLEYARRHPDYSVVLRLIAKTYETTLEKC 360  
Db 325 DLPSLAADFVSKGVKQVNAEAKDVFGLMFLEYARRHPDYSVVLRLIAKTYETTLEKC 384  
QY 361 CAAADPHECYAKVDFDFKPLVREPQNLIKONCELPQOLGEYKFNALLVRYTKVPQVST 420  
Db 385 CAAADPHECYAKVDFDFKPLVREPQNLIKONCELPQOLGEYKFNALLVRYTKVPQVST 444  
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHKTVPVSDRVTKCCTES 480  
Db 445 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHKTVPVSDRVTKCCTES 504  
QY 481 LVNRRPCFSALEVDVETVYKFNATFTFHADICTLSEKEROIKKOTALVELVGHKPKAT 540  
Db 505 LVNRRPCFSALEVDVETVYKFNATFTFHADICTLSEKEROIKKOTALVELVGHKPKAT 564  
QY 541 KEQLKAVMDDDFAAFVEKCKCADDKETCFABEGKKLVAASQAALGL- - - - -MSPRLEV 592  
Db 565 KEQLKAVMDDDFAAFVEKCKCADDKETCFABEGKKLVAASQAALGLATGLTSLDPSLEI 624  
QY 593 - - - - -PCSHALPOGL 602  
Db 625 YKMFVKKRRQQLLAKNLQALNDIHQQYKILDMVLMKGLFKVLEDSRTVLTAAADVLPDGP 684  
QY 603 SPGQVLIIVRGLVLOPKHFTVSLRDOAAHAPVTLRASPADRTLQ- - - - -W- 647  
Db 685 FP- - - - -QDEK- - - - -LKDAFSHV-VENTAFGDDVVLFRPRIVHYVFDHNSWN 727  
QY 648 -ISRMGQKKLISAPFLFYPPQRFVLELLFQEGGLK--LALNGQGLG- - - - - 690  
Db 728 LLIRWG- - - - -IS- - - - -FCNQTGVFNQGHSPHLSLMAQELGISEKDSNFQNPFKI 774  
QY 691 - - - - -ATSMNQAL- - - - -EQLREL 706  
Db 775 DRTEFIPSTDPPQKALREEEKRKRKEIR 807

RESULT 8

US-10-775-204-1614  
; Sequence 1614, Application US/10775204  
; Publication No. US20050186664A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, Craig A.  
; APPLICANT: Haseltine, William A.  
; APPLICANT: Balance, David J.  
; APPLICANT: Turner, Andrew J.  
; TITLE OF INVENTION: Albumin Fusion Proteins

FILE REFERENCE: PF564  
; CURRENT APPLICATION NUMBER: US/10/775,204  
; CURRENT FILING DATE: 2004-02-11  
; PRIOR APPLICATION NUMBER: 60/341,811  
; PRIOR FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/360,000  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: 60/378,950  
; PRIOR FILING DATE: 2002-05-10  
; PRIOR APPLICATION NUMBER: 60/398,008  
; PRIOR FILING DATE: 2002-07-24  
; PRIOR APPLICATION NUMBER: 60/411,355  
; PRIOR FILING DATE: 2002-09-18  
; PRIOR APPLICATION NUMBER: 60/414,984  
; PRIOR FILING DATE: 2002-10-02  
; PRIOR APPLICATION NUMBER: 60/417,611  
; PRIOR FILING DATE: 2002-10-11  
; PRIOR APPLICATION NUMBER: 60/420,246  
; PRIOR FILING DATE: 2002-10-23  
; PRIOR APPLICATION NUMBER: 60/423,623  
; PRIOR FILING DATE: 2002-11-05  
; PRIOR APPLICATION NUMBER: 60/351,360  
; PRIOR FILING DATE: 2002-01-28  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 2222  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 1614  
; LENGTH: 819  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-775-204-1614

Query Match 83.3%; Score 3152.3; DB 18; Length 819;  
Best Local Similarity 76.6%; Pred. No. 2.3e-75;  
Matches 623; Conservative 16; Mismatches 37; Indels 137; Gaps 12;

QY 1 DAHSEVAHFKDGLGEENFKALVLIAPQAQYLOQCPPEHVKLVNEVTEFAKTCVADSEAE 60  
DB 25 DAHSEVAHFKDGLGEENFKALVLIAPQAQYLOQCPPEHVKLVNEVTEFAKTCVADSEAE 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144  
QY 121 DVNCTAFHDNEETFLKYLVEIARRHPYFYAPPELLPEAKYKAAFTCCQAADKAAACLLP 180  
DB 145 DVNCTAFHDNEETFLKYLVEIARRHPYFYAPPELLPEAKYKAAFTCCQAADKAAACLLP 204  
QY 181 KLDELDRDEGKASAKQRLKASLQKFGERAFKAWAVARLSQRPFAEFVSKLVTDLTK 240  
DB 205 KLDELDRDEGKASAKQRLKASLQKFGERAFKAWAVARLSQRPFAEFVSKLVTDLTK 264  
QY 241 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLKCECKPPLLEKSHCHIAEVNDEMPA 300  
DB 265 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLKCECKPPLLEKSHCHIAEVNDEMPA 324  
QY 301 DLPSLAADFVSKDVCNVAEAKDVFGLGMFLYEYARHPDYSVVLLRLAKTYETTTLEKC 360  
DB 325 DLPSLAADFVSKDVCNVAEAKDVFGLGMFLYEYARHPDYSVVLLRLAKTYETTTLEKC 384  
QY 361 CAAADPHECAKVFDEKPLVEEPQNLIKNCELFEQLGEYKFNQALLVRYTKVPQVST 420  
DB 385 CAAADPHECAKVFDEKPLVEEPQNLIKNCELFEQLGEYKFNQALLVRYTKVPQVST 444  
QY 421 PTLVEVSRNLGKVGSKCKKHPKAPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 445 PTLVEVSRNLGKVGSKCKKHPKAPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 504  
QY 481 LVNRRPCFSALEVDVETVPKFNATETTFHADICTLSEKROIKKOTALVELVKKHKPAT 540  
DB 505 LVNRRPCFSALEVDVETVPKFNATETTFHADICTLSEKROIKKOTALVELVKKHKPAT 564  
QY 541 KEQLKAVMDDFAAFEKCKCADDKETCFABEGKKLVAASQAALGL-----MSPLREV 592

DB 565 KEQLKAVMDDFAAFEKCKCADDKETCFABEGKKLVAASQAALGLATGTLRTSLDSLEI 624  
QY 593 -----PCSHALPOGL 602  
DB 625 YKKMFEVKEREQLLALKNLALQNDIHQQYKILDMVKGLFKVLEDSRTVLTAADVLPDGP 684  
QY 603 SPGQVIVRGLVLOQPKHFTVSLRDOAAHAPVTLRASFADRILQ-----W- 647  
DB 685 FP-----ODEK-----LKDAFHSV-VENTAFFGVDVLRFPRIVHYYPDHNSMN 727  
QY 648 -ISRWGOKKLIAPFLFYQPRFEVLLLFQEGGLK--LALNGOGLG----- 690  
DB 728 LLIRWG-----IS-----FCNQTGVFNQGHPSILSMAQELGISEKDSNFQNPFKI 774  
QY 691 -----ATSMNQAL-----EQLRELR 706  
DB 775 DRTEFIPSTDPQKALREEEKRKKKEKEIR 807

RESULT 9  
US-10-775-204-1616  
; Sequence 1616, Application US/10775204  
; Publication No. US20050186664A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, Craig A.  
; APPLICANT: Haseltine, William A.  
; APPLICANT: Balance, David J.  
; APPLICANT: Turner, Andrew J.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF564  
; CURRENT APPLICATION NUMBER: US/10/775,204  
; CURRENT FILING DATE: 2004-02-11  
; PRIOR APPLICATION NUMBER: 60/341,811  
; PRIOR FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/360,000  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: 60/378,950  
; PRIOR FILING DATE: 2002-05-10  
; PRIOR APPLICATION NUMBER: 60/398,008  
; PRIOR FILING DATE: 2002-07-24  
; PRIOR APPLICATION NUMBER: 60/411,355  
; PRIOR FILING DATE: 2002-09-18  
; PRIOR APPLICATION NUMBER: 60/414,984  
; PRIOR FILING DATE: 2002-10-02  
; PRIOR APPLICATION NUMBER: 60/417,611  
; PRIOR FILING DATE: 2002-10-11  
; PRIOR APPLICATION NUMBER: 60/420,246  
; PRIOR FILING DATE: 2002-10-23  
; PRIOR APPLICATION NUMBER: 60/423,623  
; PRIOR FILING DATE: 2002-11-05  
; PRIOR APPLICATION NUMBER: 60/351,360  
; PRIOR FILING DATE: 2002-01-28  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 2222  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 1616  
; LENGTH: 819  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-775-204-1616

Query Match 83.3%; Score 3152.3; DB 18; Length 819;  
Best Local Similarity 76.6%; Pred. No. 2.3e-75;  
Matches 623; Conservative 16; Mismatches 37; Indels 137; Gaps 12;

QY 1 DAHSEVAHFKDGLGEENFKALVLIAPQAQYLOQCPPEHVKLVNEVTEFAKTCVADSEAE 60  
DB 25 DAHSEVAHFKDGLGEENFKALVLIAPQAQYLOQCPPEHVKLVNEVTEFAKTCVADSEAE 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144

121 DVMCTAFHNEETELKYLVEIARRHPYFVAPPELLFAKRYKAAFTCCQAAADKAACLLP 180  
145 DVMCTAFHNEETELKYLVEIARRHPYFVAPPELLFAKRYKAAFTCCQAAADKAACLLP 204  
181 KLDELDEGKASSAKORLKASLOKFGGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 240  
205 KLDELDEGKASSAKORLKASLOKFGGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 264  
241 VHTTECHGDLLECCADRADLAKYICENQDSISSKKECKEPLLEKSHCIAEVENDEMPA 300  
265 VHTTECHGDLLECCADRADLAKYICENQDSISSKKECKEPLLEKSHCIAEVENDEMPA 324  
301 DLPSLAADPVESKOVCKNYAEAKOVFLGMFLYEYARRHPDYSVLLLRLLAKTYETLEK 360  
325 DLPSLAADPVESKOVCKNYAEAKOVFLGMFLYEYARRHPDYSVLLLRLLAKTYETLEK 384  
361 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELFQELGEYKFQNALVRYTKVPQVST 420  
385 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELFQELGEYKFQNALVRYTKVPQVST 444  
421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNOLCVLHKTVPVSDRVTCKCTES 480  
445 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNOLCVLHKTVPVSDRVTCKCTES 504  
481 LVNRRPCPSALEVDETYVPKEFNAETFTPHADICTLSEKEROIKKOTALVELVKKHKKAT 540  
505 LVNRRPCPSALEVDETYVPKEFNAETFTPHADICTLSEKEROIKKOTALVELVKKHKKAT 564  
541 KEQLKAVMDDFAAFAVEKCKCADDKETCFABEGKCLVAASQAALGL-----MSPRLEV 592  
565 KEQLKAVMDDFAAFAVEKCKCADDKETCFABEGKCLVAASQAALGLATGTLRSLDPSLEI 624  
593 -----ATSMNQAL-----EQLREUR 706  
625 YKMFVKKRREQLLAKNLAQLNDIHQQYKILDVMLKGLFKVLEDSRTVLTAAADVLPDGP 684  
603 SPQVILVRLVLPKPKFTVSLRQAAHAPVTLRASADRTLQ-----W- 647  
685 FP-----QDEK-----LKDAFHSV-VENTAFFGVVLRFPRIVHYFYFDHNSWN 727  
648 -ISRWGOKKLISAPFLFYQRFEFVLLFQEGGLK--LALNQQGLG-----690  
728 LLIRWG---IS-----FCNQTGVFNQGPSPLLSLMAQELGISEKDSNFQNPFKI 774  
691 -----ATSMNQAL-----EQLREUR 706  
775 DRTEPIPTDPFOKALREBEKRRKEKRKEIR 807

RESULT 10  
US-10-775-204-1657  
; Sequence 1657, Application US/10775204  
; Publication No. US20050186664A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, Craig A.  
; APPLICANT: Haseltine, William A.  
; APPLICANT: Balance, David J.  
; APPLICANT: Turner, Andrew J.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF564  
; CURRENT APPLICATION NUMBER: US/10/775,204  
; PRIOR FILING DATE: 2004-02-11  
; PRIOR APPLICATION NUMBER: 60/341,811  
; PRIOR FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/360,000  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: 60/378,950  
; PRIOR FILING DATE: 2002-05-10  
; PRIOR APPLICATION NUMBER: 60/398,008  
; PRIOR FILING DATE: 2002-07-24  
; PRIOR APPLICATION NUMBER: 60/411,355  
; PRIOR FILING DATE: 2002-09-18

; PRIOR APPLICATION NUMBER: 60/414,984  
; PRIOR FILING DATE: 2002-10-02  
; PRIOR APPLICATION NUMBER: 60/417,611  
; PRIOR FILING DATE: 2002-10-11  
; PRIOR APPLICATION NUMBER: 60/420,246  
; PRIOR FILING DATE: 2002-10-23  
; PRIOR APPLICATION NUMBER: 60/423,623  
; PRIOR FILING DATE: 2002-11-05  
; PRIOR APPLICATION NUMBER: 60/351,360  
; PRIOR FILING DATE: 2003-01-28  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 2222  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 1657  
; LENGTH: 794  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-775-204-1657

Query Match 83.3%; Score 3152.2; DB 18; Length 794;  
Best Local Similarity 76.3%; Pred. No. 2.2e-75;  
Matches 617; Conservative 18; Mismatches 36; Indels 138; Gaps 9;

QY 1 DAHKSEVAHRPKDLGEENFKALVLIATAFOYLQCCPFEDHVKLVNVEVTEPAKTCVADESAB 60  
DB 25 DAHKSEVAHRPKDLGEENFKALVLIATAFOYLQCCPFEDHVKLVNVEVTEPAKTCVADESAB 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECFLOHKDDNPNLRLVRPEV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECFLOHKDDNPNLRLVRPEV 144  
QY 121 DVMCTAFHNEETELKYLVEIARRHPYFVAPPELLFAKRYKAAFTCCQAAADKAACLLP 180  
DB 145 DVMCTAFHNEETELKYLVEIARRHPYFVAPPELLFAKRYKAAFTCCQAAADKAACLLP 204  
QY 181 KLDELDEGKASSAKORLKASLOKFGGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 240  
DB 205 KLDELDEGKASSAKORLKASLOKFGGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 264  
QY 241 VHTTECHGDLLECCADRADLAKYICENQDSISSKKECKEPLLEKSHCIAEVENDEMPA 300  
DB 265 VHTTECHGDLLECCADRADLAKYICENQDSISSKKECKEPLLEKSHCIAEVENDEMPA 324  
QY 301 DLPSLAADPVESKOVCKNYAEAKOVFLGMFLYEYARRHPDYSVLLLRLLAKTYETLEK 360  
DB 325 DLPSLAADPVESKOVCKNYAEAKOVFLGMFLYEYARRHPDYSVLLLRLLAKTYETLEK 384  
QY 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELFQELGEYKFQNALVRYTKVPQVST 420  
DB 385 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELFQELGEYKFQNALVRYTKVPQVST 444  
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNOLCVLHKTVPVSDRVTCKCTES 480  
DB 445 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNOLCVLHKTVPVSDRVTCKCTES 504  
QY 481 LVNRRPCPSALEVDETYVPKEFNAETFTPHADICTLSEKEROIKKOTALVELVKKHKKAT 540  
DB 505 LVNRRPCPSALEVDETYVPKEFNAETFTPHADICTLSEKEROIKKOTALVELVKKHKKAT 564  
QY 541 KEQLKAVMDDFAAFAVEKCKCADDKETCFABEGKCLVAASQAALGL-----SPRLEV 592  
DB 565 KEQLKAVMDDFAAFAVEKCKCADDKETCFABEGKCLVAASQAALGLMAFTEHSPILFHRD 624  
QY 593 PCSHAL-----PQGLSPG-----605  
DB 625 LCSSRSIWLARKIRSDLTALTESYVVKHGLKNINLDSADGMBVASTDQWSEUTEAERLQE 684  
QY 606 -----QVILVRLVLPKPKFTVSLRD--QAAHAPVTLRASFA-----641  
DB 685 NLQAYRTFHVLLAR--LLEDQVQVHFPTTEGDFHQAHTHLLQLVAAAFAYQIIEMLILEYKI 743  
QY 642 -----DRTLQWISRWGQKGLISAPFLFYQRFEFVLLFQEGGLKLA 683

Db 744 PRNEADGMPINVDGGLFEKKL-----WGLKVL-----QE----- 773

Qy 684 LNGOGLGATSMNQALEQLRELRISGSVQ 712  
Db 774 -----LSQWTVRSIHDRLFSSHQ 792

## RESULT 11

US-10-775-204-1548  
; Sequence 1548, Application US/10775204  
; Publication No. US20050186664A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, Craig A.  
; APPLICANT: Haseltine, William A.  
; APPLICANT: Balance, David J.  
; APPLICANT: Turner, Andrew J.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF564  
; CURRENT APPLICATION NUMBER: US/10/775,204  
; PRIOR FILING DATE: 2004-02-11  
; PRIOR APPLICATION NUMBER: 60/341,811  
; PRIOR FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/360,000  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: 60/378,950  
; PRIOR FILING DATE: 2002-05-10  
; PRIOR APPLICATION NUMBER: 60/398,008  
; PRIOR FILING DATE: 2002-07-24  
; PRIOR APPLICATION NUMBER: 60/411,355  
; PRIOR FILING DATE: 2002-09-18  
; PRIOR APPLICATION NUMBER: 60/414,984  
; PRIOR FILING DATE: 2002-10-02  
; PRIOR APPLICATION NUMBER: 60/417,611  
; PRIOR FILING DATE: 2002-10-11  
; PRIOR APPLICATION NUMBER: 60/420,246  
; PRIOR FILING DATE: 2002-10-23  
; PRIOR APPLICATION NUMBER: 60/423,623  
; PRIOR FILING DATE: 2002-11-05  
; PRIOR APPLICATION NUMBER: 60/351,360  
; PRIOR FILING DATE: 2002-01-28  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 2222  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 1548  
; LENGTH: 809  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-775-204-1548

Query Match 83.3%; Score 3152.2; DB 18; Length 809;  
Best Local Similarity 76.3%; Pred. No. 2.2e-75;  
Matches 617; Conservative 18; Mismatches 36; Indels 138; Gaps 9;

Qy 1 DAHKSVAHRFDLGENFKALVLIAPAFYLOQCPEDHVKLVNEVTEFAKTCVADESAB 60  
Db 25 DAHKSVAHRFDLGENFKALVLIAPAFYLOQCPEDHVKLVNEVTEFAKTCVADESAB 84  
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDPNLPRVLRPEV 120  
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDPNLPRVLRPEV 144  
Qy 121 DVNCTAFHNDNEETFLKKYLYEIAARRPYFYAPPELLFFAKRYKAAPTECCOQADKAACLLP 180  
Db 145 DVNCTAFHNDNEETFLKKYLYEIAARRPYFYAPPELLFFAKRYKAAPTECCOQADKAACLLP 204  
Qy 181 KLDELDEGKASSAKQRLKCSLOKFGERAFKAWAVARLSQRPFAEFVSKLVTDLTK 240  
Db 205 KLDELDEGKASSAKQRLKCSLOKFGERAFKAWAVARLSQRPFAEFVSKLVTDLTK 264  
Qy 241 VHTECCHGDLLECADRADLAKYICENODSISKKLKECKECPKLEKSHCICAEVDEMPA 300  
Db 265 VHTECCHGDLLECADRADLAKYICENQDSISKKLKECKECPKLEKSHCICAEVDEMPA 324

Qy 301 DLPSLAADPVSKDKYKNAEAKVDYFLGMYLYEYARRHPDYVWLLLRLLAKTYETTLK 360  
Db 325 DLPSLAADPVSKDKYKNAEAKVDYFLGMYLYEYARRHPDYVWLLLRLLAKTYETTLK 384  
Qy 361 CAADAPHECYAKVDFEKPPLVEEPQNLKQNCLEFQELGEYKFNALLVRYTKVPQVST 420  
Db 385 CAADAPHECYAKVDFEKPPLVEEPQNLKQNCLEFQELGEYKFNALLVRYTKVPQVST 444  
Qy 421 PTLVEVSRNLGKVGSKCCGHPKAPCAEDYLSVNLQLCVHLHEKTPVSDRVTCKCTES 480  
Db 445 PTLVEVSRNLGKVGSKCCGHPKAPCAEDYLSVNLQLCVHLHEKTPVSDRVTCKCTES 504  
Qy 481 LVNRRPCPSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVKKHKPKAT 540  
Db 505 LVNRRPCPSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVKKHKPKAT 564  
Qy 541 KEOLKAVNDDFAAFVEKCKKADDDKTCFAEEGKGLVAASQAALGLM-----SPRLEV 592  
Db 565 KEOLKAVNDDFAAFVEKCKKADDDKTCFAEEGKGLVAASQAALGLM-----SPRLEV 624  
Qy 593 PCSHAL-----PQGLSPG----- 605  
Db 625 LCSRSIWLARKIRSOLTALTESYVVKHQLNKNINLDSADGMPVASTDQWSELTEARLQE 684  
Qy 606 -----QVIIVRGVLQEPKHFVTSLED--QAHAHAPVTLRASA----- 641  
Db 685 NLQAYRTFHVLLAR--LLEDQQVHFPTTEGDFHQAHTLLQLQAAVAQYEEMLILEYKI 743  
Qy 642 -----DRTLQWISRWGQKKLISAPFLFYQRPFEVLLLFQEGGLKLA 683  
Db 744 PRNEADGMPINVDGGLFEKKL-----WGLKVL-----QE----- 773  
Qy 684 LNGOGLGATSMNQALEQLRELRISGSVQ 712  
Db 774 -----LSQWTVRSIHDRLFSSHQ 792

## RESULT 12

US-10-775-204-364  
; Sequence 364, Application US/10775204  
; Publication No. US20050186664A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, Craig A.  
; APPLICANT: Haseltine, William A.  
; APPLICANT: Balance, David J.  
; APPLICANT: Turner, Andrew J.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF564  
; CURRENT APPLICATION NUMBER: US/10/775,204  
; CURRENT FILING DATE: 2004-02-11  
; PRIOR APPLICATION NUMBER: 60/341,811  
; PRIOR FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/360,000  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: 60/378,950  
; PRIOR FILING DATE: 2002-05-10  
; PRIOR APPLICATION NUMBER: 60/398,008  
; PRIOR FILING DATE: 2002-07-24  
; PRIOR APPLICATION NUMBER: 60/411,355  
; PRIOR FILING DATE: 2002-09-18  
; PRIOR APPLICATION NUMBER: 60/414,984  
; PRIOR FILING DATE: 2002-10-02  
; PRIOR APPLICATION NUMBER: 60/417,611  
; PRIOR FILING DATE: 2002-10-11  
; PRIOR APPLICATION NUMBER: 60/420,246  
; PRIOR FILING DATE: 2002-10-23  
; PRIOR APPLICATION NUMBER: 60/423,623  
; PRIOR FILING DATE: 2002-11-05  
; PRIOR APPLICATION NUMBER: 60/351,360  
; PRIOR FILING DATE: 2002-01-28  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 2222  
; SOFTWARE: Patentin Ver. 2.0

```
; SEQ ID NO 364
; LENGTH: 774
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-775-204-364

Query Match      83.3%; Score 3152; DB 18; Length 774;
Best Local Similarity 80.3%; Pred. No. 2.1e-75;
Matches 615; Conservative 10; Mismatches 31; Indels 110; Gaps 7;

QY 1 DAHSEVAHRFKDLGEENFKALVLIAPFAQYIQQCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 25 DAHSEVAHRFKDLGEENFKALVLIAPFAQYIQQCPFEDHVKLVNEVTEFAKTCVADESAAE 84
QY 61 NCDKSLHHTLFGDKLCTVATLTRETYGEMADCCAKQEPERNECFLOHKDNDPNLPLRVREV 120
DB 85 NCDKSLHHTLFGDKLCTVATLTRETYGEMADCCAKQEPERNECFLOHKDNDPNLPLRVREV 144
QY 121 DVMTAFHDNEETFLKYLIEIARRHPYFYAPPELLFFAKRYKAAAFTECCQAADKAACLLP 180
DB 145 DVMTAFHDNEETFLKYLIEIARRHPYFYAPPELLFFAKRYKAAAFTECCQAADKAACLLP 204
QY 181 KLDELRDGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQRFPAKAEFAVSKLVTDLTK 240
DB 205 KLDELRDGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQRFPAKAEFAVSKLVTDLTK 264
QY 241 VHTCCHGDLLECADDDRADLAKYICENQDSISSKLKECCPKPFLLEKSHCIAEVENDENPA 300
DB 265 VHTCCHGDLLECADDDRADLAKYICENQDSISSKLKECCPKPFLLEKSHCIAEVENDENPA 324
QY 301 DLPSLAADFVSKDKVCKNYAEAKOVFLGMFLYEVARRHPDYSVLLLRKLTAKTYETTLK 360
DB 325 DLPSLAADFVSKDKVCKNYAEAKOVFLGMFLYEVARRHPDYSVLLLRKLTAKTYETTLK 384
QY 361 CAADPHCEYAKVDFEFPKPLVEEPQNLIKONCELFQELGEYKFNQALLVRYTKVPQVST 420
DB 385 CAADPHCEYAKVDFEFPKPLVEEPQNLIKONCELFQELGEYKFNQALLVRYTKVPQVST 444
QY 421 PTLVEVSRLGKVGSKCKHPKAEKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 445 PTLVEVSRLGKVGSKCKHPKAEKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDVETYPVKPFNAETFTFHADICTLSEKEROIKKQOTALVELVKKPKAT 540
DB 505 LVNRRPCFSALEVDVETYPVKPFNAETFTFHADICTLSEKEROIKKQOTALVELVKKPKAT 564
QY 541 KEQLKAVMDDFAAFVEKCKKADDKTCFAEBGKKLVAASQAALGLMSPLR----- 590
DB 565 KEQLKAVMDDFAAFVEKCKKADDKTCFAEBGKKLVAASQAALGLAPPRLLICDSRVLERY 624
QY 591 -----EVPQSHA-----LPOGLS 603
DB 625 LLEAKEAEQITTCGAHCSLNEQITVPDPTKVNFYAWKRMVEVQQQAVEVWQGLALLSEAVL 684
QY 604 PQGVILVRLGLVLEP-----KHFTVSLRDQAHAHPVTLSAFSADRTLQWISRW 651
DB 685 RQGLALVQSSQPWEPLQLHVDKAVSGLSLTLTLR-----ALRA----- 723
QY 652 GQKKLISAP-----FLFYPQRFPEVLLLFQEGGLKL 682
DB 724 -QKEAISPPDAASAAPLRTITADTF---RKLFRVYSNFLRGKCLKL 764
```

## RESULT 13

```
US-10-775-204-366
; Sequence 366, Application US/10775204
; Publication No. US20050186664A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; APPLICANT: Balance, David J.
; APPLICANT: Turner, Andrew J.
; TITLE OF INVENTION: Albumin Fusion Proteins
```

```
; FILE REFERENCE: PF564
; CURRENT APPLICATION NUMBER: US/10/775,204
; CURRENT FILING DATE: 2004-02-11
; PRIOR APPLICATION NUMBER: 60/341,811
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/360,000
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/378,950
; PRIOR FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: 60/398,008
; PRIOR FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: 60/411,355
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/414,984
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: 60/417,611
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: 60/420,246
; PRIOR FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: 60/423,623
; PRIOR FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: 60/351,360
; PRIOR FILING DATE: 2002-01-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2222
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 366
; LENGTH: 774
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-775-204-366
```

```
Query Match      83.3%; Score 3152; DB 18; Length 774;
Best Local Similarity 80.3%; Pred. No. 2.1e-75;
Matches 615; Conservative 10; Mismatches 31; Indels 110; Gaps 7;
```

```
QY 1 DAHSEVAHRFKDLGEENFKALVLIAPFAQYIQQCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 25 DAHSEVAHRFKDLGEENFKALVLIAPFAQYIQQCPFEDHVKLVNEVTEFAKTCVADESAAE 84
QY 61 NCDKSLHHTLFGDKLCTVATLTRETYGEMADCCAKQEPERNECFLOHKDNDPNLPLRVREV 120
DB 85 NCDKSLHHTLFGDKLCTVATLTRETYGEMADCCAKQEPERNECFLOHKDNDPNLPLRVREV 144
QY 121 DVMTAFHDNEETFLKYLIEIARRHPYFYAPPELLFFAKRYKAAAFTECCQAADKAACLLP 180
DB 145 DVMTAFHDNEETFLKYLIEIARRHPYFYAPPELLFFAKRYKAAAFTECCQAADKAACLLP 204
QY 181 KLDELRDGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQRFPAKAEFAVSKLVTDLTK 240
DB 205 KLDELRDGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQRFPAKAEFAVSKLVTDLTK 264
QY 241 VHTCCHGDLLECADDDRADLAKYICENQDSISSKLKECCPKPFLLEKSHCIAEVENDENPA 300
DB 265 VHTCCHGDLLECADDDRADLAKYICENQDSISSKLKECCPKPFLLEKSHCIAEVENDENPA 324
QY 301 DLPSLAADFVSKDKVCKNYAEAKOVFLGMFLYEVARRHPDYSVLLLRKLTAKTYETTLK 360
DB 325 DLPSLAADFVSKDKVCKNYAEAKOVFLGMFLYEVARRHPDYSVLLLRKLTAKTYETTLK 384
QY 361 CAADPHCEYAKVDFEFPKPLVEEPQNLIKONCELFQELGEYKFNQALLVRYTKVPQVST 420
DB 385 CAADPHCEYAKVDFEFPKPLVEEPQNLIKONCELFQELGEYKFNQALLVRYTKVPQVST 444
QY 421 PTLVEVSRLGKVGSKCKHPKAEKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 445 PTLVEVSRLGKVGSKCKHPKAEKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDVETYPVKPFNAETFTFHADICTLSEKEROIKKQOTALVELVKKPKAT 540
DB 505 LVNRRPCFSALEVDVETYPVKPFNAETFTFHADICTLSEKEROIKKQOTALVELVKKPKAT 564
QY 541 KEQLKAVMDDFAAFVEKCKKADDKTCFAEBGKKLVAASQAALGLMSPLR----- 590
```

```
Db 565 KEQLKAVNMDFAAFVEKCKKADDDKTCFAEGKGLVAASQAALGLAPRLICDSRVLERY 624
Qy 591 -----EVPCSHA-----LPOGLS 603
Db 625 LLEAKEARQITTCGAHCSLNEQITVPDTKVNIFYAKMEVCOQAVVWQGLALSEAVL 684
Qy 604 PQQVIIRGLVLEQEP-----KHFTVSLRDQAHAHAPVTLRASFAADRTLQWISRW 651
Db 685 RQOALLVQSSQPWEPLQLHVDKAVSGLSLTLLR-----ALRA-----723
Qy 652 GQKLIISAP-----FLFYQRFVLEVLLFOEGGLKL 682
Db 724 -QKEAISPPDAASAAPLRTITADTF-----RKLFRVYSNFLRGKGL 764

RESULT 14
US-10-775-204-245
; Sequence 245, Application US/10775204
; Publication No. US20050186664A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; APPLICANT: Balance, David J.
; APPLICANT: Turner, Andrew J.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF564
; CURRENT APPLICATION NUMBER: US/10/775,204
; CURRENT FILING DATE: 2004-02-11
; PRIOR APPLICATION NUMBER: 60/341,811
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/360,000
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/378,950
; PRIOR FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: 60/398,008
; PRIOR FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: 60/411,355
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/414,984
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: 60/417,611
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: 60/420,246
; PRIOR FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: 60/423,623
; PRIOR FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: 60/351,360
; PRIOR FILING DATE: 2002-01-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 222
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 245
; LENGTH: 774
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-775-204-245

Query Match 83.2%; Score 3151; DB 18; Length 774;
Best Local Similarity 80.3%; Pred. No. 2.2e-75;
Matches 615; Conservative 9; Mismatches 32; Indels 110; Gaps 7;

Qy 1 DAHKSVAHRFDLGENFKALVLIAPQVLOQCPEDHVKLVNEVTEFAKTCVADESAE 60
Db 25 DAHKSVAHRFDLGENFKALVLIAPQVLOQCPEDHVKLVNEVTEFAKTCVADESAE 84
Qy 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Db 85 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144
Qy 121 DVNCTAFHNEETFLKKYLYEIARRHPYFAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180
Db 145 DVNCTAFHNEETFLKKYLYEIARRHPYFAPPELLFFAKRYKAAFTTECCQAADKAACLLP 204
```

```
Qy 181 KDELRLDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQRFPPKAEFAEVSCLVTDLT 240
Db 205 KDELRLDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQRFPPKAEFAEVSCLVTDLT 264
Qy 241 VHTECHGDLLLECCADDDRADLAKYICENODSISLKLKECCPKLLEKSHCIAEVENDEMPA 300
Db 265 VHTECHGDLLLECCADDDRADLAKYICENODSISLKLKECCPKLLEKSHCIAEVENDEMPA 324
Qy 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVVLRLAKTYETTLK 360
Db 325 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVVLRLAKTYETTLK 384
Qy 361 CAADPHCYAKVDFEFPPLVEEPONLIKONCELFEQLGEYKFQVALLVRYTKKVPQVST 420
Db 385 CAADPHCYAKVDFEFPPLVEEPONLIKONCELFEQLGEYKFQVALLVRYTKKVPQVST 444
Qy 421 PTLVEVSRNLGKVGSKCKCKHPEAKMPCAEYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 445 PTLVEVSRNLGKVGSKCKCKHPEAKMPCAEYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504
Qy 481 LVNRRPCFSALEVDYVVPKEFNAETFTFHADICTLSEKERQIKQTALVELVGHKPKAT 540
Db 505 LVNRRPCFSALEVDYVVPKEFNAETFTFHADICTLSEKERQIKQTALVELVGHKPKAT 564
Qy 541 KEQLKAVNMDFAAFVEKCKKADDDKTCFAEGKGLVAASQAALGLMSPL-----590
Db 565 KEQLKAVNMDFAAFVEKCKKADDDKTCFAEGKGLVAASQAALGLAPRLICDSRVLERY 624
Qy 591 -----EVPCSHA-----LPOGLS 603
Db 625 LLEAKEARENITTCGAHCSLNEITVPDTKVNIFYAKMEVCOQAVVWQGLALSEAVL 684
Qy 604 PQQVIIRGLVLEQEP-----KHFTVSLRDQAHAHAPVTLRASFAADRTLQWISRW 651
Db 685 RQOALLVQSSQPWEPLQLHVDKAVSGLSLTLLR-----ALRA-----723
Qy 652 GQKLIISAP-----FLFYQRFVLEVLLFOEGGLKL 682
Db 724 -QKEAISPPDAASAAPLRTITADTF-----RKLFRVYSNFLRGKGL 764

RESULT 15
US-10-775-204-255
; Sequence 255, Application US/10775204
; Publication No. US20050186664A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; APPLICANT: Balance, David J.
; APPLICANT: Turner, Andrew J.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF564
; CURRENT APPLICATION NUMBER: US/10/775,204
; CURRENT FILING DATE: 2004-02-11
; PRIOR APPLICATION NUMBER: 60/341,811
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/360,000
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/378,950
; PRIOR FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: 60/398,008
; PRIOR FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: 60/411,355
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/414,984
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: 60/417,611
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: 60/420,246
; PRIOR FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: 60/423,623
; PRIOR FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: 60/351,360
; PRIOR FILING DATE: 2002-01-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 222
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 245
; LENGTH: 774
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-775-204-255
```

```

; PRIOR APPLICATION NUMBER: 60/351,360
; PRIOR FILING DATE: 2002-01-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 222
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 255
; LENGTH: 774
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-775-204-255

Query Match      83.2%; Score 3151; DB 18; Length 774;
Best Local Similarity 80.5%; Pred. No. 2.2e-75;
Matches 617; Conservative 9; Mismatches 30; Indels 110; Gaps 8;

Qy 1 DAKHSEVAHAFKDLGEENFKALVLIAPAYLQOCPPFDHVKLVNEVTEFAKTCVADSSAE 60
Db 25 DAKHSEVAHAFKDLGEENFKALVLIAPAYLQOCPPFDHVKLVNEVTEFAKTCVADSSAE 84
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPMLPLRVPEV 120
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPMLPLRVPEV 144
Qy 121 DVMCTAFHDNBEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180
Db 145 DVMCTAFHDNBEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 204
Qy 181 KLDELDEGKASSAKORLKCASLOKFGERAFAKAVARLSORFPKAPFAEVSKLVTDLTK 240
Db 205 KLDELDEGKASSAKORLKCASLOKFGERAFAKAVARLSORFPKAPFAEVSKLVTDLTK 264
Qy 241 VHTCECHGDLLECADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 300
Db 265 VHTCECHGDLLECADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 324
Qy 301 DLPSLAADPVESKDVCKNYAEAKDVFLGMFLYFYEARHPDYSVVLRLAKTYETTLK 360
Db 325 DLPSLAADPVESKDVCKNYAEAKDVFLGMFLYFYEARHPDYSVVLRLAKTYETTLK 384
Qy 361 CAAADPHECYAKVDFEKPPLVEEPQNLIKONCELLFEQGLGEYKFNALLVRYTKKVPQVST 420
Db 385 CAAADPHECYAKVDFEKPPLVEEPQNLIKONCELLFEQGLGEYKFNALLVRYTKKVPQVST 444
Qy 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 445 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504
Qy 481 LVNRRPCFSALVEDETYYVPKEFNATFTFHADICTLSEKEROIKKQTALVELVKHKPKAT 540
Db 505 LVNRRPCFSALVEDETYYVPKEFNATFTFHADICTLSEKEROIKKQTALVELVKHKPKAT 564
Qy 541 KEQLKAVMDDFAAFVEKCKCKADDDKTCFAERGGKLVAAASQAALGLMSPR----- 589
Db 565 KEQLKAVMDDFAAFVEKCKCKADDDKTCFAERGGKLVAAASQAALGLAPPLRCDISRVLE 624
Qy 590 -LEVPCKSHALP-----OGLS----- 603
Db 625 LLEAKEBAEITTCBAEHCSLNEATVPTDKNFYANKMEVGGQAVVWQGLALLSEAVL 684
Qy 604 PGQVITVIRGLVLOEP-----KHFTVSLRDQAAHAPVTLRASFAADRTLQWISRW 651
Db 685 RGQALLVASSQPWEPLQLHVDKAVSGLSLTLRL-----ALRA----- 723
Qy 652 GQKKLISAP-----FLFYPQRFVEVILLFQEGGLKL 682
Db 724 -QKEAISPPDAASAAPLFTITADTF-----RKLFRVYSNFLRGKGL 764

RESULT 16
US-10-775-204-256
; Sequence 256, Application US/10775204
; Publication No. US2005018664A1
; GENERAL INFORMATION:

```





;; PRIOR FILING DATE: 2002-10-11  
;; PRIOR APPLICATION NUMBER: 60/420,246  
;; PRIOR FILING DATE: 2002-10-23  
;; PRIOR APPLICATION NUMBER: 60/423,623  
;; PRIOR FILING DATE: 2002-11-05  
;; PRIOR APPLICATION NUMBER: 60/351,360  
;; PRIOR FILING DATE: 2002-01-28  
;; Remaining Prior Application data removed - See File Wrapper or PALM.  
;; NUMBER OF SEQ ID NOS: 222  
;; SOFTWARE: Patentin Ver. 2.0  
;; SEQ ID NO 302  
;; LENGTH: 774  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-775-204-302

Query Match 83.2%; Score 3151; DB 18; Length 774;  
Best Local Similarity 80.3%; Pred. No. 2.2e-75;  
Matches 615; Conservative 9; Mismatches 32; Indels 110; Gaps 7;

Qy 1 DAHSEVAHRPKDLGEENFKALVLIAPAYLQOCPFEDHVKLVNVEVTEFAKTCVADSSAE 60  
Db 25 DAHSEVAHRPKDLGEENFKALVLIAPAYLQOCPFEDHVKLVNVEVTEFAKTCVADSSAE 84

Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNNPRLVPRV 120  
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNNPRLVPRV 144

Qy 121 DVMCTAFHNDNETFLKKLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
Db 145 DVMCTAFHNDNETFLKKLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP 204

Qy 181 KLDELDRGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRFPPKAEFAEVSCLVTDLT 240  
Db 205 KLDELDRGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRFPPKAEFAEVSCLVTDLT 264

Qy 241 VHTCCGGDLLECCADDDRADLAKYICENQDSISSKLEKCEKPELLEKSHCIAEVENDEMPA 300  
Db 265 VHTCCGGDLLECCADDDRADLAKYICENQDSISSKLEKCEKPELLEKSHCIAEVENDEMPA 324

Qy 301 DLPSLAADFVBSKDVCKNYAEAKDVFLGMFLYIARRHPDYSVVLLRLAKTYETTLTK 360  
Db 325 DLPSLAADFVBSKDVCKNYAEAKDVFLGMFLYIARRHPDYSVVLLRLAKTYETTLTK 384

Qy 361 CAADPHCEYAKVDFEFPKLVVEEPQNLIKQNCFLQGEYKFNQALLVRYTKKVPQVST 420  
Db 385 CAADPHCEYAKVDFEFPKLVVEEPQNLIKQNCFLQGEYKFNQALLVRYTKKVPQVST 444

Qy 421 PTLVEVSRLGKVGSKCKKHPKAMPKAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
Db 445 PTLVEVSRLGKVGSKCKKHPKAMPKAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504

Qy 481 LVNRRPCFSALEVDYTPKGFNAETFTFHADICTLSEKERQIKKQATLALVELVKHKPKAT 540  
Db 505 LVNRRPCFSALEVDYTPKGFNAETFTFHADICTLSEKERQIKKQATLALVELVKHKPKAT 564

Qy 541 KEQLKAVMDDDFAAFVEKCKADDKETCPAEBGKLVAAASQAALGMSRPL----- 590  
Db 565 KEQLKAVMDDDFAAFVEKCKADDKETCPAEBGKLVAAASQAALGMSRPL----- 624

Qy 591 -----EVPCHSA-----LPQGLS 603  
Db 625 LLEAKEAENITTGCAEHCSLNNITVPTDKVNFYAWKRMVEVQQQAVEVMQGLALLSEAVL 684

Qy 604 PGQVLIIVRGLVLQEP-----KHFTVSLRDQAHAHPVTLRASFADRTLQWISRW 651  
Db 685 RQALLVNSQWEPQLHVDKAVSGLSLTLTLR-----ALRA----- 723

Qy 652 GQKKLISAP-----FLFYPRFFEVALLFOEGGLKL 682  
Db 724 -QKEAISPDAAASAAPLRTITADTF- ---RKLFRVYSNFLRGKGLKL 764

RESULT 19  
US-10-775-180-72  
; Sequence 72, Application US/10775180  
; Publication No. US20050054570A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, Craig A.  
; APPLICANT: Haseltine, William A.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PFS74  
; CURRENT APPLICATION NUMBER: US/10775,180  
; CURRENT FILING DATE: 2004-02-11  
; PRIOR APPLICATION NUMBER: PCT/US02/40892  
; PRIOR FILING DATE: 2002-12-23  
; PRIOR APPLICATION NUMBER: 60/341,811  
; PRIOR FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/360,000  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: 60/378,950  
; PRIOR FILING DATE: 2002-05-10  
; PRIOR APPLICATION NUMBER: 60/398,008  
; PRIOR FILING DATE: 2002-07-24  
; PRIOR APPLICATION NUMBER: 60/411,355  
; PRIOR FILING DATE: 2002-09-18  
; PRIOR APPLICATION NUMBER: 60/414,984  
; PRIOR FILING DATE: 2002-10-02  
; PRIOR APPLICATION NUMBER: 60/417,611  
; PRIOR FILING DATE: 2002-10-11  
; PRIOR APPLICATION NUMBER: 60/420,246  
; PRIOR FILING DATE: 2002-10-23  
; PRIOR APPLICATION NUMBER: 60/423,623  
; PRIOR FILING DATE: 2002-11-05  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 858  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 72  
; LENGTH: 868  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-775-180-72

Query Match 83.2%; Score 3151; DB 17; Length 868;  
Best Local Similarity 70.5%; Pred. No. 2.7e-75;  
Matches 621; Conservative 17; Mismatches 33; Indels 210; Gaps 11;

Qy 1 DAHSEVAHRPKDLGEENFKALVLIAPAYLQOCPFEDHVKLVNVEVTEFAKTCVADSSAE 60  
Db 25 DAHSEVAHRPKDLGEENFKALVLIAPAYLQOCPFEDHVKLVNVEVTEFAKTCVADSSAE 84

Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNNPRLVPRV 120  
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNNPRLVPRV 144

Qy 121 DVMCTAFHNDNETFLKKLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
Db 145 DVMCTAFHNDNETFLKKLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP 204

Qy 181 KLDELDRGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRFPPKAEFAEVSCLVTDLT 240  
Db 205 KLDELDRGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRFPPKAEFAEVSCLVTDLT 264

Qy 241 VHTCCGGDLLECCADDDRADLAKYICENQDSISSKLEKCEKPELLEKSHCIAEVENDEMPA 300  
Db 265 VHTCCGGDLLECCADDDRADLAKYICENQDSISSKLEKCEKPELLEKSHCIAEVENDEMPA 324

Qy 301 DLPSLAADFVBSKDVCKNYAEAKDVFLGMFLYIARRHPDYSVVLLRLAKTYETTLTK 360  
Db 325 DLPSLAADFVBSKDVCKNYAEAKDVFLGMFLYIARRHPDYSVVLLRLAKTYETTLTK 384

Qy 361 CAADPHCEYAKVDFEFPKLVVEEPQNLIKQNCFLQGEYKFNQALLVRYTKKVPQVST 420  
Db 385 CAADPHCEYAKVDFEFPKLVVEEPQNLIKQNCFLQGEYKFNQALLVRYTKKVPQVST 444

Qy 421 PTLVEVSRLGKVGSKCKKHPKAMPKAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480

```
Db 445 PTLVEVSRLGKVGSKCKHPKAEKMPKCAEDYLSVNLQCLVHKEKTPVSDRVTKCCTES 504
Qy 481 LVNRRPCFSALEVDYVYVPEKFNATFTFFHADICTLSEKERQIKKQATLVELVGHKPKAT 540
Db 505 LVNRRPCFSALEVDYVYVPEKFNATFTFFHADICTLSEKERQIKKQATLVELVGHKPKAT 564
Qy 541 KEQLKAVNMDFAAFVEKCKKADDETCFAEBGKGLVAASQAALGLMSPLRLEVP 593
Db 565 KEQLKAVNMDFAAFVEKCKKADDETCFAEBGKGLVAASQAALGLVA 621
Qy 594 -----CSHALPQGLSPGVIIVRGLVLE-----PKHFT----- 622
Db 622 ETGERLVCAQCPP-----GTFVQPCRRDSPTTCGCPPRHYTFWNYLRCR 669
Qy 623 -----VSLRDQAHA----- 632
Db 670 YCNVLCGEREBEARACHATHNRACRCRTGFFAHAGFCLHASCPCPGAGVIAPGTPSQNTQ 729
Qy 633 -----PVTLRAS----- 639
Db 730 COPCPGTFSSSSSSSEQOCPHRNCTALGLALNVPSSSHDTLCTSGTGFPLSTRVPGA 789
Qy 640 -----FADRTLQWISR-----WG-----QKKLISAPFLFYPPORF 668
Db 790 ECERAVIDFVAFQDISIKRLQRLQALQALQALQALQALQALQALQALQALQALQALQAL 840
Qy 669 FEVLFFQEGGLKALNGOGLGATSMNQOALEBQLRELKISG 709
Db 841 LTELGAQDQALLVRL-----LQALRVARMPG 867
```

## RESULT 20

```
US-10-775-204-219
; Sequence 219, Application US/10775204
; Publication No. US20050186664A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; APPLICANT: Balance, David J.
; APPLICANT: Turner, Andrew J.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF564
; CURRENT APPLICATION NUMBER: US/10/775,204
; CURRENT FILING DATE: 2004-02-11
; PRIOR APPLICATION NUMBER: 60/341,811
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/360,000
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/378,950
; PRIOR FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: 60/398,008
; PRIOR FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: 60/411,355
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/414,984
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: 60/417,611
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: 60/420,246
; PRIOR FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: 60/423,623
; PRIOR FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: 60/351,360
; PRIOR FILING DATE: 2002-01-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2222
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 219
; LENGTH: 868
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-775-204-219
```

```
Query Match 83.2%; Score 3151; DB 18; Length 868;
Best Local Similarity 70.5%; Pred. No. 2.7e-75;
Matches 621; Conservative 17; Mismatches 33; Indels 210; Gaps 11;
```

```
Qy 1 DAHSEVAHRFKDLGEENFKALVLIAPAOYLQOCPPEFDHVKLVNVEVTEFAKTCVADESAAE 60
Db 25 DAHSEVAHRFKDLGEENFKALVLIAPAOYLQOCPPEFDHVKLVNVEVTEFAKTCVADESAAE 84
Qy 61 NCDLSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLRVREV 120
Db 85 NCDLSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLRVREV 144
Qy 121 DVMCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAAFTECCOAAADKAACLLP 180
Db 145 DVMCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAAFTECCOAAADKAACLLP 204
Qy 181 KLDELREGKASSAKORLKASLOKFGGERAFKAWARLSORFPKAEFAEYSKLVTDLT 240
Db 205 KLDELREGKASSAKORLKASLOKFGGERAFKAWARLSORFPKAEFAEYSKLVTDLT 264
Qy 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 300
Db 265 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 324
Qy 301 DLPSLAADFVESKDVCKNYABAKOVFLGMFLYEVARRHPDYSVVLLLLLAKTYETTLK 360
Db 325 DLPSLAADFVESKDVCKNYABAKOVFLGMFLYEVARRHPDYSVVLLLLLAKTYETTLK 384
Qy 361 CAADPHCEYAKVDFEKPFLVEEPQNLIKONCELFQELGEYKFNQALVRYTKVPQVST 420
Db 385 CAADPHCEYAKVDFEKPFLVEEPQNLIKONCELFQELGEYKFNQALVRYTKVPQVST 444
Qy 421 PTLVEVSRLGKVGSKCKHPKAEKMPKCAEDYLSVNLQCLVHKEKTPVSDRVTKCCTES 480
Db 445 PTLVEVSRLGKVGSKCKHPKAEKMPKCAEDYLSVNLQCLVHKEKTPVSDRVTKCCTES 504
Qy 481 LVNRRPCFSALEVDYVYVPEKFNATFTFFHADICTLSEKERQIKKQATLVELVGHKPKAT 540
Db 505 LVNRRPCFSALEVDYVYVPEKFNATFTFFHADICTLSEKERQIKKQATLVELVGHKPKAT 564
Qy 541 KEQLKAVNMDFAAFVEKCKKADDETCFAEBGKGLVAASQAALGLMSPLRLEVP 593
Db 565 KEQLKAVNMDFAAFVEKCKKADDETCFAEBGKGLVAASQAALGLVA 621
Qy 594 -----CSHALPQGLSPGVIIVRGLVLE-----PKHFT----- 622
Db 622 ETGERLVCAQCPP-----GTFVQPCRRDSPTTCGCPPRHYTFWNYLRCR 669
Qy 623 -----VSLRDQAHA----- 632
Db 670 YCNVLCGEREBEARACHATHNRACRCRTGFFAHAGFCLHASCPCPGAGVIAPGTPSQNTQ 729
Qy 633 -----PVTLRAS----- 639
Db 730 COPCPGTFSSSSSSSEQOCPHRNCTALGLALNVPSSSHDTLCTSGTGFPLSTRVPGA 789
Qy 640 -----FADRTLQWISR-----WG-----QKKLISAPFLFYPPORF 668
Db 790 ECERAVIDFVAFQDISIKRLQRLQALQALQALQALQALQALQALQALQALQALQALQAL 840
Qy 669 FEVLFFQEGGLKALNGOGLGATSMNQOALEBQLRELKISG 709
Db 841 LTELGAQDQALLVRL-----LQALRVARMPG 867
```

## RESULT 21

```
US-10-775-180-71
; Sequence 71, Application US/10775180
; Publication No. US20050054570A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
```



Db 25 DAHSEVAHRFKDLGEENFKALVLIAPQYLOQCFFEDHVKLVNEVTEFAKTCVADESAAE 84  
Qy 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVREV 120  
Db 85 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVREV 144  
Qy 121 DVMCTAFHNEBTFLLKYLIEIARHPYFAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180  
Db 145 DVMCTAFHNEBTFLLKYLIEIARHPYFAPPELLFFAKRYKAAFTTECCQAADKAACLLP 204  
Qy 181 KLDELDEGKASSAKORLKCSLQKGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240  
Db 205 KLDELDEGKASSAKORLKCSLQKGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 264  
Qy 241 VHTCCHGDLLECADRADLAKYICENQDSISSKLKECEKPELLEKSHCIAEVENDEMPA 300  
Db 265 VHTCCHGDLLECADRADLAKYICENQDSISSKLKECEKPELLEKSHCIAEVENDEMPA 324  
Qy 301 DLPSLAADFVESKOVCKNYAEAKDVFGLMFLYIYARRHPDYVSVLLRLAKTYETTLK 360  
Db 325 DLPSLAADFVESKOVCKNYAEAKDVFGLMFLYIYARRHPDYVSVLLRLAKTYETTLK 384  
Qy 361 CAADPHECYAKVDFEFLVEEPONLIKONCELFQOLGEYKFQNALVRYTKVPQVST 420  
Db 385 CAADPHECYAKVDFEFLVEEPONLIKONCELFQOLGEYKFQNALVRYTKVPQVST 444  
Qy 421 PTLVEVRNLGKVGSKCKKHPEAKMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCTES 480  
Db 445 PTLVEVRNLGKVGSKCKKHPEAKMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCTES 504  
Qy 481 LVNRRPCFSALEVDVETYPKFNATFTFPHADICTLSEKERQIKQTALVELVKGHPKAT 540  
Db 505 LVNRRPCFSALEVDVETYPKFNATFTFPHADICTLSEKERQIKQTALVELVKGHPKAT 564  
Qy 541 KEQLKAVMDDFAAFEKCKKADDKETCFABEGKLVAAASQAALGLVA---ETPTVWRDA 593  
Db 565 KEQLKAVMDDFAAFEKCKKADDKETCFABEGKLVAAASQAALGLVA---ETPTVWRDA 621  
Qy 594 -----CSHALPQGLSPQGVIIIVRGLVLOE-----PKHFT----- 622  
Db 622 ETGERLVCAQCPP-----GTFVQPCRRDSPTTCGCPFRHVTQFWNYLERCR 669  
Qy 623 -----VSLRDQAHA----- 632  
Db 670 YCNVLGGEREEBARACHATHNACRTRTGFAHAGFLEHASCPPGAGVIAPGTPSONTO 729  
Qy 633 -----PVTLRAS----- 639  
Db 730 CQCPPEGTFSSSSSSSQCPHNRNCTALGLALNVPGSSSHDTLCTSGTFPLSTRVPGAE 789  
Qy 640 -----FADRTLQWISR-----WG-----QKLLISAPFLFVPQRF 668  
Db 790 ECERAVIDFVAQDISIKELQRLQLALEAPEGNGPTPRAGRAALQLK-----RRR 840  
Qy 669 FEVLLLFQBGGLKIALNGOGLGATSNNOQALSOQLRELISG 709  
Db 841 LTELIGAOQDGAALLVRL-----LQALRVARMFG 867

RESULT 23  
US-10-775-180-70  
; Sequence 70, Application US/10775180  
; Publication No. US20050054570A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, Craig A.  
; APPLICANT: Haseltine, William A.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF574  
; CURRENT APPLICATION NUMBER: US/10/775,180  
; CURRENT FILING DATE: 2004-02-11  
; PRIOR APPLICATION NUMBER: PCT/US02/40892  
; PRIOR FILING DATE: 2002-12-23  
; PRIOR APPLICATION NUMBER: 60/341,811

; PRIOR FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/360,000  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: 60/378,950  
; PRIOR FILING DATE: 2002-05-10  
; PRIOR APPLICATION NUMBER: 60/398,008  
; PRIOR FILING DATE: 2002-07-24  
; PRIOR APPLICATION NUMBER: 60/411,355  
; PRIOR FILING DATE: 2002-09-18  
; PRIOR APPLICATION NUMBER: 60/414,984  
; PRIOR FILING DATE: 2002-10-02  
; PRIOR APPLICATION NUMBER: 60/417,611  
; PRIOR FILING DATE: 2002-10-11  
; PRIOR APPLICATION NUMBER: 60/420,246  
; PRIOR FILING DATE: 2002-10-23  
; PRIOR APPLICATION NUMBER: 60/423,623  
; PRIOR FILING DATE: 2002-11-05  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 858  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 70  
; LENGTH: 880  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-775-180-70

Query Match 83.2%; Score 3151; DB 17; Length 880;  
Best Local Similarity 70.5%; Pred. No. 2.8e-75;  
Matches 621; Conservative 17; Mismatches 33; Indels 210; Gaps 11;

Qy 1 DAHSEVAHRFKDLGEENFKALVLIAPQYLOQCFFEDHVKLVNEVTEFAKTCVADESAAE 60  
Db 25 DAHSEVAHRFKDLGEENFKALVLIAPQYLOQCFFEDHVKLVNEVTEFAKTCVADESAAE 84  
Qy 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVREV 120  
Db 85 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVREV 144  
Qy 121 DVMCTAFHNEBTFLLKYLIEIARHPYFAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180  
Db 145 DVMCTAFHNEBTFLLKYLIEIARHPYFAPPELLFFAKRYKAAFTTECCQAADKAACLLP 204  
Qy 181 KLDELDEGKASSAKORLKCSLQKGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240  
Db 205 KLDELDEGKASSAKORLKCSLQKGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 264  
Qy 241 VHTCCHGDLLECADRADLAKYICENQDSISSKLKECEKPELLEKSHCIAEVENDEMPA 300  
Db 265 VHTCCHGDLLECADRADLAKYICENQDSISSKLKECEKPELLEKSHCIAEVENDEMPA 324  
Qy 301 DLPSLAADFVESKOVCKNYAEAKDVFGLMFLYIYARRHPDYVSVLLRLAKTYETTLK 360  
Db 325 DLPSLAADFVESKOVCKNYAEAKDVFGLMFLYIYARRHPDYVSVLLRLAKTYETTLK 384  
Qy 361 CAADPHECYAKVDFEFLVEEPONLIKONCELFQOLGEYKFQNALVRYTKVPQVST 420  
Db 385 CAADPHECYAKVDFEFLVEEPONLIKONCELFQOLGEYKFQNALVRYTKVPQVST 444  
Qy 421 PTLVEVRNLGKVGSKCKKHPEAKMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCTES 480  
Db 445 PTLVEVRNLGKVGSKCKKHPEAKMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCTES 504  
Qy 481 LVNRRPCFSALEVDVETYPKFNATFTFPHADICTLSEKERQIKQTALVELVKGHPKAT 540  
Db 505 LVNRRPCFSALEVDVETYPKFNATFTFPHADICTLSEKERQIKQTALVELVKGHPKAT 564  
Qy 541 KEQLKAVMDDFAAFEKCKKADDKETCFABEGKLVAAASQAALGLVA---ETPTVWRDA 593  
Db 565 KEQLKAVMDDFAAFEKCKKADDKETCFABEGKLVAAASQAALGLVA---ETPTVWRDA 621  
Qy 594 -----CSHALPQGLSPQGVIIIVRGLVLOE-----PKHFT----- 622  
Db 622 ETGERLVCAQCPP-----GTFVQPCRRDSPTTCGCPFRHVTQFWNYLERCR 669

Qy 623 -----VSLRDOAHNA----- 632  
Db 670 YCNVLCGEREEARACHATHNRACRGTGFFAHAGFCLEHASCPGAGVIAPGTPSONTO 729  
Qy 633 -----PVTLRAS----- 639  
Db 730 QCPCPGTFSSSSSSSQCPHNRCTALGLALNVPGSSSHDTLCTSGTGPLSTRVPGAE 789  
Qy 640 -----FADRTLQWISR-----WG-----OKKLISAPFLFYPORF 668  
Db 790 ECERAVIDFVAFQDISIKRLQLLQALEAPEGWGPTPRAGRAALQLKL-----RRR 840  
Qy 669 FEVLILFQEGGLKALNGQGLGATSMNQQALEQLRELIRISG 709  
Db 841 LTELGLAQDQDGLLVRL-----LQALRVARMMPG 867

RESULT 24

US-10-775-180-74  
; Sequence 74, Application US/10775180  
; Publication No. US20050054570A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, Craig A.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF574  
; CURRENT FILING DATE: 2004-02-11  
; PRIOR APPLICATION NUMBER: PCT/US02/40892  
; PRIOR FILING DATE: 2002-12-23  
; PRIOR APPLICATION NUMBER: 60/341,811  
; PRIOR FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/360,000  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: 60/378,950  
; PRIOR FILING DATE: 2002-05-10  
; PRIOR APPLICATION NUMBER: 60/398,008  
; PRIOR FILING DATE: 2002-07-24  
; PRIOR APPLICATION NUMBER: 60/411,355  
; PRIOR FILING DATE: 2002-09-18  
; PRIOR APPLICATION NUMBER: 60/414,984  
; PRIOR FILING DATE: 2002-10-02  
; PRIOR APPLICATION NUMBER: 60/417,611  
; PRIOR FILING DATE: 2002-10-11  
; PRIOR APPLICATION NUMBER: 60/420,246  
; PRIOR FILING DATE: 2002-10-23  
; PRIOR APPLICATION NUMBER: 60/423,623  
; PRIOR FILING DATE: 2002-11-05  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 858  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 74  
; LENGTH: 880  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-775-180-74

Query Match 83.2%; Score 3151; DB 17; Length 880;  
Best Local Similarity 70.5%; Pred. No. 2.8e-75;  
Matches 621; Conservative 17; Mismatches 33; Indels 210; Gaps 11;  
Qy 1 DAKHSEVAHRFKDLGEENFKALVLAFAQYLOQCFFEDHVKLVNEVTEFAKTCVADESAE 60  
Db 25 DAKHSEVAHRFKDLGEENFKALVLAFAQYLOQCFFEDHVKLVNEVTEFAKTCVADESAE 84  
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLPRLVREPV 120  
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLPRLVREPV 144  
Qy 121 DVMCTAFHNDNEFTFKKLYETARRHPYFAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180  
Db 145 DVMCTAFHNDNEFTFKKLYETARRHPYFAPPELLFFAKRYKAAFTCCQAAADKAACLLP 204

Qy 181 KLDELDEGKASSAKORLKCASLOKFGERAFAKAWAVARLSORFPPKAEFAEVSKLVDLT 240  
Db 205 KLDELDEGKASSAKORLKCASLOKFGERAFAKAWAVARLSORFPPKAEFAEVSKLVDLT 264  
Qy 241 VHTECHGDILLECADDRAIDLAKYICENQDISISKLKECCCKPPLLEKSHCHIAEVENDEMPA 300  
Db 265 VHTECHGDILLECADDRAIDLAKYICENQDISISKLKECCCKPPLLEKSHCHIAEVENDEMPA 324  
Qy 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYIYARRHPDYSVLLLRLAKTETTTLEKC 360  
Db 325 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYIYARRHPDYSVLLLRLAKTETTTLEKC 384  
Qy 361 CAAADPHECVAKYVDFEPLVEBPONLIKONCELFEOLGEYKFTQNALLVRYTKKVPVOST 420  
Db 385 CAAADPHECVAKYVDFEPLVEBPONLIKONCELFEOLGEYKFTQNALLVRYTKKVPVOST 444  
Qy 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480  
Db 445 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 504  
Qy 481 LVNRRPCFSALEVDVETVPKFEFNAETFTFHADICTLSEKEROIKKOTALVELVKHKEPAT 540  
Db 505 LVNRRPCFSALEVDVETVPKFEFNAETFTFHADICTLSEKEROIKKOTALVELVKHKEPAT 564  
Qy 541 KEQLKAVMDDFAAFVEKCKKADDETCFABEGKKLVAASQAALGLMSPRLEVP----- 593  
Db 565 KEQLKAVMDDFAAFVEKCKKADDETCFABEGKKLVAASQAALGLVA---ETPTYWMDA 621  
Qy 594 -----CSHALPQGLSPGVIIIVRGLVLOE-----PKHFT----- 622  
Db 622 ETGERLVCAQCPP-----GTFVQRPCHRDSPPTTCGCPCHRYTQFWMYLERCR 669  
Qy 623 -----VSLRDOAHNA----- 632  
Db 670 YCNVLCGEREEARACHATHNRACRGTGFFAHAGFCLEHASCPGAGVIAPGTPSONTO 729  
Qy 633 -----PVTLRAS----- 639  
Db 730 QCPCPGTFSSSSSSSQCPHNRCTALGLALNVPGSSSHDTLCTSGTGPLSTRVPGAE 789  
Qy 640 -----FADRTLQWISR-----WG-----OKKLISAPFLFYPORF 668  
Db 790 ECERAVIDFVAFQDISIKRLQLLQALEAPEGWGPTPRAGRAALQLKL-----RRR 840  
Qy 669 FEVLILFQEGGLKALNGQGLGATSMNQQALEQLRELIRISG 709  
Db 841 LTELGLAQDQDGLLVRL-----LQALRVARMMPG 867

RESULT 25

US-10-775-180-79  
; Sequence 79, Application US/10775180  
; Publication No. US20050054570A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, Craig A.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF574  
; CURRENT FILING DATE: 2004-02-11  
; PRIOR APPLICATION NUMBER: PCT/US02/40892  
; PRIOR FILING DATE: 2002-12-23  
; PRIOR APPLICATION NUMBER: 60/341,811  
; PRIOR FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/360,000  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: 60/378,950  
; PRIOR FILING DATE: 2002-05-10  
; PRIOR APPLICATION NUMBER: 60/398,008  
; PRIOR FILING DATE: 2002-07-24  
; PRIOR APPLICATION NUMBER: 60/411,355  
; PRIOR FILING DATE: 2002-09-18

QY 640 -----FADBTLOWISR-----WG-----QKKLISAPFLVFPORF 668  
DB 790 ECERAVIDFVAFQDISIKRQLORLQALBAPEGWGTTPRAGRALQQLK-----RRR 840  
QY 669 FEVILLFOEGGLKLAINGQGLGATSMNQQALEQLRELRIISG 709  
DB 841 LTELGAQDQALLVRL-----LQALVARWPG 867

RESULT 26  
US-10-775-180-87  
; Sequence 87, Application US/10775180  
; Publication No. US20050054570A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, Craig A.  
; APPLICANT: Haseltine, William A.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF574  
; CURRENT APPLICATION NUMBER: US/10775,180  
; PRIORITY FILING DATE: 2004-02-11  
; PRIOR APPLICATION NUMBER: PCT/US02/40892  
; PRIOR FILING DATE: 2002-12-23  
; PRIOR APPLICATION NUMBER: 60/341,811  
; PRIOR FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/360,000  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: 60/378,950  
; PRIOR FILING DATE: 2002-05-10  
; PRIOR APPLICATION NUMBER: 60/398,008  
; PRIOR FILING DATE: 2002-07-24  
; PRIOR APPLICATION NUMBER: 60/411,355  
; PRIOR FILING DATE: 2002-09-18  
; PRIOR APPLICATION NUMBER: 60/414,984  
; PRIOR FILING DATE: 2002-10-02  
; PRIOR APPLICATION NUMBER: 60/417,611  
; PRIOR FILING DATE: 2002-10-11  
; PRIOR APPLICATION NUMBER: 60/420,246  
; PRIOR FILING DATE: 2002-10-23  
; PRIOR APPLICATION NUMBER: 60/423,623  
; PRIOR FILING DATE: 2002-11-05  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 858  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 87  
; LENGTH: 880  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-775-180-87

Query Match 83.2%; Score 3151; DB 17; Length 880;  
Best Local Similarity 70.5%; Pred. No. 2.8e-75;  
Matches 621; Conservative 17; Mismatches 33; Indels 210; Gaps 11;

QY 1 DAHSEVAHRFDKLGSENFKALVLAFAQYLOQCPEDHVKLVNEVTEFAKTCVADESAE 60  
DB 25 DAHSEVAHRFDKLGSENFKALVLAFAQYLOQCPEDHVKLVNEVTEFAKTCVADESAE 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVLRPEV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVLRPEV 144  
QY 121 DVMTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
DB 145 DVMTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 204  
QY 181 KLDELDEGKASSAKORLKASLOKFGERAFKAWAVARLSORPPKAEFAEVSCLVTDLT 240  
DB 205 KLDELDEGKASSAKORLKASLOKFGERAFKAWAVARLSORPPKAEFAEVSCLVTDLT 264  
QY 241 VHTCCGHGDLLECADDRADLAKYICENODSISSKLKECCCKPILLESKSHCIAEVENDEMPA 300  
DB 265 VHTCCGHGDLLECADDRADLAKYICENODSISSKLKECCCKPILLESKSHCIAEVENDEMPA 324

QY 1 DAHSEVAHRFDKLGSENFKALVLAFAQYLOQCPEDHVKLVNEVTEFAKTCVADESAE 60  
DB 25 DAHSEVAHRFDKLGSENFKALVLAFAQYLOQCPEDHVKLVNEVTEFAKTCVADESAE 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVLRPEV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVLRPEV 144  
QY 121 DVMTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
DB 145 DVMTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 204  
QY 181 KLDELDEGKASSAKORLKASLOKFGERAFKAWAVARLSORPPKAEFAEVSCLVTDLT 240  
DB 205 KLDELDEGKASSAKORLKASLOKFGERAFKAWAVARLSORPPKAEFAEVSCLVTDLT 264  
QY 241 VHTCCGHGDLLECADDRADLAKYICENODSISSKLKECCCKPILLESKSHCIAEVENDEMPA 300  
DB 265 VHTCCGHGDLLECADDRADLAKYICENODSISSKLKECCCKPILLESKSHCIAEVENDEMPA 324  
QY 301 DPLSLAADFVSKDVCNVAEAKDVLGMFLFYIARRHPDYSVLLLRLLAKYIETTLK 360  
DB 325 DPLSLAADFVSKDVCNVAEAKDVLGMFLFYIARRHPDYSVLLLRLLAKYIETTLK 384  
QY 361 CAAADPHECVAKYVDFEKLVEPQNLIKONCELFEQLGEYKFNALLVRYTKKVPQVST 420  
DB 385 CAAADPHECVAKYVDFEKLVEPQNLIKONCELFEQLGEYKFNALLVRYTKKVPQVST 444  
QY 421 PTLVEVSRNLGKVGSKCCXHPKRMPCAEEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 445 PTLVEVSRNLGKVGSKCCXHPKRMPCAEEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504  
QY 481 LVNRRFCFSALEVDETVVPKEFNAETFTTHADICTLSEKERQIKKOTALVELVGHKPKAT 540  
DB 505 LVNRRFCFSALEVDETVVPKEFNAETFTTHADICTLSEKERQIKKOTALVELVGHKPKAT 564  
QY 541 KEOLKAVMDDFAAVFEKCKKADDDKTCFAEKGKLVAAQALGLMSPLRLEP----- 593  
DB 565 KEOLKAVMDDFAAVFEKCKKADDDKTCFAEKGKLVAAQALGLVA---ETPTYWRDA 621  
QY 594 -----CShALPQGLSPGQVIVRGLVQ-----PKHFT----- 622  
DB 622 ETGERLVCAQCP-----GTFVORPCRDSPPTTCGCPPRHYTFQFNWYLERCR 669  
QY 623 -----VSLRQQAHA----- 632  
DB 670 YCNVLGEREEERACHATHNRACRGTGFFAHAGFLEHASCPPGAGVIAPCTPSONTQ 729  
QY 633 -----PVTLRAS----- 639  
DB 730 CQCPPTGTFSSASSSSSEQCPHQRCTALGLALNVPGSSSHDTLCTSGTGPLSTRVPGAE 789



```
QY 301 DLPSLAADPVESKDVCKNYAEAKOVFLGMFLYEVARRHPDYSVLLLELAKTYETTLK 360
D 325 DLPSLAADPVESKDVCKNYAEAKOVFLGMFLYEVARRHPDYSVLLLELAKTYETTLK 384
QY 361 CAADPHCEYAKVDFEPKPLVEEONLQKNCFLGEQYKFNQALLVRYTKVPQVST 420
D 385 CAADPHCEYAKVDFEPKPLVEEONLQKNCFLGEQYKFNQALLVRYTKVPQVST 444
QY 421 PTLVEVSRLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
D 445 PTLVEVSRLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDETYVPKFNATFTFHADICTLSEKERQIKKQATLVELVHKPKAT 540
D 505 LVNRRPCFSALEVDETYVPKFNATFTFHADICTLSEKERQIKKQATLVELVHKPKAT 564
QY 541 KEQLKAVMDDFAAFVEKCKKADDDKTCFAEBGKKLVAASQAALGLMSPRLEVP 593
D 565 KEQLKAVMDDFAAFVEKCKKADDDKTCFAEBGKKLVAASQAALGLMSPRLEVP 621
QY 594 -----CShALPOGLSPGVIIIRGLVLQE-----PKHFT----- 622
D 622 ETGERLVCAQCQP-----GTFVQPCRRDSTTCGCPPRHVTQFWNYLRCR 669
QY 670 YCNVLCGEREBEARACHATHNRCRCRTGTFFAHAGFCLHASCPPGAGVIAPGTPSQNTQ 729
D 633 -----PVTLRAS----- 639
QY 730 COPCPPTGFSASSSSSQCPHRNCTALGLALNVPSSSHDTLCTSCGFPPLSTVRPGAE 789
D 640 -----FADRTLQWISR-----WG-----QKKLISAPFLVPORF 668
QY 790 ECERAVIDFAVQDISIKRLQRLQALEAPEGWGPTPRAGRAALQLKL-----RRR 840
D 669 FEVLLFQEGGLKALNGQGLGATSMNOQALEQLRELISG 709
D 841 LTELGAQDQALLVRL-----LQALRVARMPG 867
```

## RESULT 27

```
US-10-775-180-90
; Sequence 90, Application US/10775180
; Publication No. US20050054570A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PFS74
; CURRENT APPLICATION NUMBER: US/10/775,180
; CURRENT FILING DATE: 2004-02-11
; PRIOR APPLICATION NUMBER: PCT/US02/40892
; PRIOR FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: 60/341,811
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/360,000
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/378,950
; PRIOR FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: 60/398,008
; PRIOR FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: 60/411,355
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/414,984
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: 60/417,611
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: 60/420,246
; PRIOR FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: 60/423,623
; PRIOR FILING DATE: 2002-11-05
; Remaining Prior Application data removed - See File Wrapper or PALM.
```

```
; NUMBER OF SEQ ID NOS: 858
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 90
; LENGTH: 880
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-775-180-90

Query Match      83.2%; Score 3151; DB 17; Length 880;
Best Local Similarity 70.5%; Pred. No. 2.8e-75;
Matches 621; Conservative 17; Mismatches 33; Indels 210; Gaps 11;

QY 1 DAHSEVAHRRPKDAGEENFKALVLIARFAQYIQQCPFFEDHVKLVNVEVTEFAKTCVADESAAE 60
D 25 DAHSEVAHRRPKDAGEENFKALVLIARFAQYIQQCPFFEDHVKLVNVEVTEFAKTCVADESAAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
D 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144
QY 121 DVMCTAFHDNBEETFLKKLYBIARRHPYFYAPPELLFFAKRYKAAATECCOAAADKAACLLP 180
D 145 DVMCTAFHDNBEETFLKKLYBIARRHPYFYAPPELLFFAKRYKAAATECCOAAADKAACLLP 204
QY 181 KLDELIRDEGKASSAKORLKASLOKFGGERAFKAWAVARLSORFPKAEPAEYVKLVTLDTK 240
D 205 KLDELIRDEGKASSAKORLKASLOKFGGERAFKAWAVARLSORFPKAEPAEYVKLVTLDTK 264
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 300
D 265 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADPVESKDVCKNYAEAKOVFLGMFLYEVARRHPDYSVLLLELAKTYETTLK 360
D 325 DLPSLAADPVESKDVCKNYAEAKOVFLGMFLYEVARRHPDYSVLLLELAKTYETTLK 384
QY 361 CAADPHCEYAKVDFEPKPLVEEONLQKNCFLGEQYKFNQALLVRYTKVPQVST 420
D 385 CAADPHCEYAKVDFEPKPLVEEONLQKNCFLGEQYKFNQALLVRYTKVPQVST 444
QY 421 PTLVEVSRLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
D 445 PTLVEVSRLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDETYVPKFNATFTFHADICTLSEKERQIKKQATLVELVHKPKAT 540
D 505 LVNRRPCFSALEVDETYVPKFNATFTFHADICTLSEKERQIKKQATLVELVHKPKAT 564
QY 541 KEQLKAVMDDFAAFVEKCKKADDDKTCFAEBGKKLVAASQAALGLMSPRLEVP 593
D 565 KEQLKAVMDDFAAFVEKCKKADDDKTCFAEBGKKLVAASQAALGLMSPRLEVP 621
QY 594 -----CShALPOGLSPGVIIIRGLVLQE-----PKHFT----- 622
D 622 ETGERLVCAQCQP-----GTFVQPCRRDSTTCGCPPRHVTQFWNYLRCR 669
QY 670 YCNVLCGEREBEARACHATHNRCRCRTGTFFAHAGFCLHASCPPGAGVIAPGTPSQNTQ 729
D 633 -----PVTLRAS----- 639
QY 730 COPCPPTGFSASSSSSQCPHRNCTALGLALNVPSSSHDTLCTSCGFPPLSTVRPGAE 789
D 640 -----FADRTLQWISR-----WG-----QKKLISAPFLVPORF 668
QY 790 ECERAVIDFAVQDISIKRLQRLQALEAPEGWGPTPRAGRAALQLKL-----RRR 840
D 669 FEVLLFQEGGLKALNGQGLGATSMNOQALEQLRELISG 709
D 841 LTELGAQDQALLVRL-----LQALRVARMPG 867
```

```

RESULT 28
US-10-775-204-217
; Sequence 217, Application US/10775204
; Publication No. US20050186664A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; APPLICANT: Balance, David J.
; APPLICANT: Turner, Andrew J.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF564
; CURRENT APPLICATION NUMBER: US/10/775,204
; CURRENT FILING DATE: 2004-02-11
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/341,811
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/360,000
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/378,950
; PRIOR FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: 60/398,008
; PRIOR FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: 60/411,355
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/414,984
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: 60/417,611
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: 60/420,246
; PRIOR FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: 60/423,623
; PRIOR FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: 60/351,360
; PRIOR FILING DATE: 2002-01-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2222
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 217
; LENGTH: 880
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-775-204-217

Query Match      83.2%; Score 3151; DB 18; Length 880;
Best Local Similarity 70.5%; Pred. No. 2.8e-75;
Matches 621; Conservative 17; Mismatches 33; Indels 210; Gaps 11;

QY 1 DAHKEVAHRFDKLGNEENFKALVLTAFAYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60
DB 25 DAHKEVAHRFDKLGNEENFKALVLTAFAYLQCCPFEDHVKLVNEVTEFAKTCVADESAB 84

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQPERNECFLOKHDDNPNLRLVRPRV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQPERNECFLOKHDDNPNLRLVRPRV 144

QY 121 DVMCTAFHDNEETFLKKLYEYIARRHPYYPAPPELLFFAKRYKAAFTCECCQAADKAACLLP 180
DB 145 DVMCTAFHDNEETFLKKLYEYIARRHPYYPAPPELLFFAKRYKAAFTCECCQAADKAACLLP 204

QY 181 KLDELDEGKASAKQRLKCSLOKQGERAFKAWAVARLSQRPKAEFAEVSRLVTDLTG 240
DB 205 KLDELDEGKASAKQRLKCSLOKQGERAFKAWAVARLSQRPKAEFAEVSRLVTDLTG 264

QY 241 VHTCECHGDLLECCADRADLAKYICENQDISISKLKECCCKPPLLEKSHCIAEVENDEMPA 300
DB 265 VHTCECHGDLLECCADRADLAKYICENQDISISKLKECCCKPPLLEKSHCIAEVENDEMPA 324

QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYEYARRHPDYSVLLRLAKTYTETLEKC 360
DB 325 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYEYARRHPDYSVLLRLAKTYTETLEKC 384

QY 361 CAAADPHECVAKVDFEFKPLVEEPQNLIIKQNCFLPEQLGEYKFNALLVRYTKKVPQVST 420
DB 385 CAAADPHECVAKVDFEFKPLVEEPQNLIIKQNCFLPEQLGEYKFNALLVRYTKKVPQVST 444

421 PTLVEVSRNLGKVGSKCKKHPBAKMPCAEDYLSVWLNQLCVLHEKTPVSDRVTKCCTES 480
445 PTLVEVSRNLGKVGSKCKKHPBAKMPCAEDYLSVWLNQLCVLHEKTPVSDRVTKCCTES 504
481 LVNRRPCFSALEVEDETVVPKEFNAETFTFHADICTLSEKERQIKQTALVELVKKPKAT 540
505 LVNRRPCFSALEVEDETVVPKEFNAETFTFHADICTLSEKERQIKQTALVELVKKPKAT 564
541 KEQLKAVMDDFAAFVEKCKKADDETCFAEEGKKLVAAASQAALGMSRLEVP----- 593
565 KEQLKAVMDDFAAFVEKCKKADDETCFAEEGKKLVAAASQAALGLVA---ETPTYPWRDA 621
594 -----CSHALPOGLSPGVIIVRGLVLOF-----PKHFT----- 622
622 ETGERLVCAQCPP-----GTFVQRPCCRRDSTTCGCPCPRHRYTOFWNYLERCR 669
623 -----VSLRDQAAHA----- 632
670 YCNVLCGEREEERACHATHNRACRCRTGTFPAHAGFCLEHASCPGAGVIAPGTSPQNTQ 729
633 -----PVTLRAS----- 639
730 CQCPPGTGFSSSSSSSEOCQPHRNCTALGLALNVPGSSSHDTLCTCTGFPPLSTRVPGA 789
640 -----PADRTLOWISR-----WG-----QKLLISAPFLFYPQRF 668
790 ECERAVIDFVAFQDISIKRLQLQALBAPEGWGTTPRAGRAALQLKL-----RRR 840
669 FEVLLFOEGGLKALNGQGLGATSMNQALAEQLELRISG 709
841 LTELIGAQDGALLVRL-----LQALVARMPG 867

RESULT 29
US-10-775-204-221
; Sequence 221, Application US/10775204
; Publication No. US20050186664A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; APPLICANT: Balance, David J.
; APPLICANT: Turner, Andrew J.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF564
; CURRENT APPLICATION NUMBER: US/10/775,204
; CURRENT FILING DATE: 2004-02-11
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/341,811
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/360,000
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/378,950
; PRIOR FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: 60/398,008
; PRIOR FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: 60/411,355
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/414,984
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: 60/417,611
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: 60/420,246
; PRIOR FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: 60/423,623
; PRIOR FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: 60/351,360
; PRIOR FILING DATE: 2002-01-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2222
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 221
; LENGTH: 880
; TYPE: PRT
```



**us-10-933-523-18.rapb**

[illegible]

Search completed: October 17, 2005, 08:35:48  
Job time : 180 secs

**THIS PAGE BLANK (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2005, 08:25:57 ; Search time 44 Seconds  
(without alignments)  
1570.083 Million cell updates/sec

Title: US-10-933-523-18  
Perfect score: 3785  
Sequence: 1 DAHSEVAHRFKDLGEENFK.....leqireirsgsvqlycvhs 718

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.1

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 100 summaries

Database : PIR\_79:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3103	82.0	609	1 ABHUS	serum albumin prec
2	2942	77.7	600	2 A47391	serum albumin prec
3	2820	69.2	608	2 S57632	serum albumin prec
4	2475.9	65.4	607	1 ABHOS	serum albumin prec
5	2446.9	64.6	607	1 ABHOS	serum albumin prec
6	2432.9	64.3	607	1 ABHOS	serum albumin prec
7	2426	64.1	608	1 ABHOS	serum albumin prec
8	2416.8	63.9	605	1 ABPGS	serum albumin prec
9	2387	63.1	609	2 JCS838	albumin - Mongolia
10	1861	49.2	453	2 A05139	serum albumin - mo
11	1557.9	41.2	615	1 ABCHS	serum albumin prec
12	1256.7	33.2	609	2 JCA258	alpha-fetoprotein
13	1252.7	33.1	609	1 FPHU	alpha-fetoprotein
14	1245.7	32.9	609	1 FPGO	alpha-fetoprotein
15	1206.6	31.9	607	1 ABXL72	74K albumin prec
16	1181.9	31.2	265	2 I46986	68K serum albumin
17	1178.3	31.1	608	1 ABXL68	alpha-fetoprotein
18	1086.8	28.7	605	1 FPM	alpha-fetoprotein
19	1071.6	28.3	611	1 FPPT	74K albumin prec
20	1059	28.0	599	1 A54906	afamin precursor -
21	933.4	24.7	614	2 S59517	serum albumin prec
22	932	24.6	608	2 A53195	afamin precursor
23	765.3	20.2	608	1 ABONS1	serum albumin 1 pr
24	761.3	20.1	608	1 ABONS2	serum albumin 2 pr
25	700.4	18.5	382	2 A37253	serum albumin - bu
26	531.4	14.0	1423	1 S27941	serum albumin - se
27	421	11.1	476	1 VYRTD	vitamin D-binding
28	417.5	11.0	474	1 VYHUD	vitamin D-binding
29	414	10.9	472	1 A35327	vitamin D-binding

30	332.1	8.8	3796	2	T18514	lysosomal traffic
31	329.6	8.7	1927	2	G64585	cag pathogenicity
32	327.2	8.6	1819	2	A71928	cag island protein
33	324	8.6	2154	2	A84669	hypothetical prote
34	322.8	8.5	2429	1	SJHUA	spectrin alpha cha
35	321.6	8.5	3951	1	VFIHB1	F1 protein - avian
36	320.5	8.5	1920	2	A53188	parietrin - mous
37	318.2	8.4	3006	2	T28625	variant-specific s
38	317.6	8.4	2210	1	RRXPTV	genome polyprotein
39	317.2	8.4	2168	2	T30171	ninein - mouse
40	316	8.3	4872	2	S27272	ryanodine receptor
41	313.4	8.3	4063	2	T42993	probable spectrin
42	313.4	8.3	4101	2	T23630	hypothetical prote
43	312.2	8.2	26926	1	I38344	titin, cardiac mus
44	311.7	8.2	4868	2	B54161	ryanodine-binding
45	310	8.2	3051	2	S42373	hypothetical prote
46	309.5	8.2	1303	2	E96805	polyketide synthas
47	309.5	8.2	4273	2	C69679	hypothetical prote
48	308.4	8.1	1529	2	T20986	nuclear migration
49	308.4	8.1	2748	2	S57976	lysosomal traffic
50	307.5	8.1	3788	2	T30851	ataxia telangiecta
51	307.4	8.1	2514	2	T37320	ryanodine receptor
52	306.8	8.1	4969	2	A37113	nuclear/mitotic ap
53	306.5	8.1	2253	2	T30336	hypothetical prote
54	306.5	8.1	2619	2	T24588	guanine nucleotide
55	306.2	8.1	2488	2	T42739	hemocytin - silkw
56	305.9	8.1	3133	2	S52093	calcium-binding pr
57	304.6	8.0	1560	2	T30282	elastic titin - hu
58	304.6	8.0	7962	2	I38346	genome polyprotein
59	304.1	8.0	2210	1	RRXPLC	protein T6D22.14
60	304	8.0	2254	2	D86215	cation-independent
61	303	8.0	2470	2	I50726	neurofibromatosis-
62	302	8.0	2818	2	B55282	neurofibromin I -
63	301.4	8.0	2820	2	JCS196	polyketide synthas
64	300.1	7.9	4447	2	A69679	type I fatty acid
65	300.1	7.9	8243	2	T31307	major merozoite su
66	299.7	7.9	1772	2	A45532	telomerase-associ
67	299.7	7.9	2629	2	T30987	protein kinase f6c
68	298.4	7.9	4128	2	JC6306	mprp protein - mo
69	298.1	7.9	1979	2	JW0059	dynein beta heavy
70	298.1	7.9	4466	1	S17231	giantin - human
71	297.8	7.9	3225	2	I52300	giantin - human
72	297.8	7.9	3259	1	A56539	insulin-like growt
73	297.5	7.9	2491	1	A28372	hypothetical prote
74	297.3	7.9	3672	2	T23433	probable laminin a
75	297.3	7.9	3704	2	T37316	bacitracin synthet
76	297.2	7.9	2607	2	T31678	genome polyprotein
77	297.1	7.8	3068	1	A44062	hypothetical prote
78	296.9	7.8	1896	2	T01490	genome polyprotein
79	296.8	7.8	2269	1	JQ1750	polyketide synthas
80	296.8	7.8	2478	2	AH2140	364K Golgi complex
81	296.7	7.8	3187	2	JCS837	variant-specific s
82	296.5	7.8	2706	2	T28155	variant-specific s
83	296.4	7.8	1932	2	T25525	variant-specific s
84	296.3	7.8	1711	2	C71625	breast cancer tumo
85	296.3	7.8	3418	1	G02334	breast cancer tumo
86	296.1	7.8	3329	2	T42205	myosin heavy chain
87	296.1	7.8	3329	2	T30904	clathrin heavy cha
88	295.9	7.8	1515	2	S51824	beige protein homo
89	295.9	7.8	2477	1	SJCHA	beige protein homo
90	295.6	7.8	1640	2	T09522	gramicidin S synth
91	295.2	7.8	3788	2	T13960	variant surface pr
92	295	7.8	4450	2	JX0340	protein F21H11.2
93	294.8	7.8	3026	2	T28431	neurofibromin - mo
94	294.7	7.8	2700	2	D88450	utrophin - human
95	294.7	7.8	2825	2	I54352	hypothetical prote
96	294.6	7.8	3433	1	S28381	genome polyprotein
97	294.5	7.8	1780	2	T17272	genome polyprotein
98	294.5	7.8	2105	1	A44059	genome polyprotein
99	294.5	7.8	4488	1	RIHM2	protein P2K11.14
100	294.4	7.8	1518	2	D96660	protein P2K11.14

Fused





R;He, X.M.; Carter, D.C.  
 Nature 358, 209-215, 1992  
 A>Title: Atomic structure and chemistry of human serum albumin.  
 A'Reference number: A46756; MUID:92334427; PMID:1630489  
 A'Contents: annotation; X-ray crystallography, 2.8 angstroms  
 R;Brown, J.R.; Shockey, P.; Behrens, P.Q.  
 In The Chemistry and Physiology of the Human Plasma Proteins, Bing, D.H., ed., pp.23-40,  
 A'Reference number: A94442  
 A'Contents: annotation; three-dimensional structure and disulfide bonds  
 R;Saber, M.A.; Stockbauer, P.; Moravsek, L.; Meloun, B.  
 Collect. Czech. Chem. Commun. 42, 564-579, 1977  
 A>Title: Disulfide bonds in human serum albumin.  
 A'Reference number: A90930  
 A'Contents: annotation; disulfide bonds  
 R;Jacobsen, C.  
 Biochem. J. 171, 453-459, 1978  
 A>Title: Lysine residue 240 of human serum albumin is involved in high-affinity binding  
 A'Reference number: A90299; MUID:78186830; PMID:656055  
 A'Contents: annotation; bilirubin-binding site  
 R;Peters, T.; Reed, R.G.  
 in Albumin: Structure, Biosynthesis, Function, Peters, J., and Sjolholm, I., eds., 11-20,  
 A>Title: Serum albumin: conformation and active sites.  
 A'Reference number: A94408  
 A'Contents: annotation; binding sites  
 R;Harper, M.E.; Dugaiczky, A.  
 Am. J. Hum. Genet. 35, 565-572, 1983  
 A>Title: Linkage of the evolutionarily-related serum albumin and alpha-fetoprotein genes  
 A'Reference number: A90028; MUID:83279982; PMID:6192711  
 A'Contents: annotation; gene position  
 R;Walker, J.E.  
 FEBS Lett. 66, 173-175, 1976  
 A>Title: Lysine residue 199 of human serum albumin is modified by acetylsalicylic acid.  
 A'Reference number: A46755; MUID:76257808; PMID:955075  
 A'Contents: annotation  
 A>Note: the nonenzymatic transfer of an acetyl group from aspirin (acetylsalicylic acid)  
 R;Bohney, J.P.; Fonda, M.L.; Feldhoff, R.C.  
 FEBS Lett. 298, 266-268, 1992  
 A>Title: Identification of Lys(190) as the primary binding site for pyridoxal 5'-phosphate  
 A'Reference number: A56294; MUID:92183881; PMID:1544460  
 A'Contents: annotation  
 A>Note: the nonenzymatic binding of pyridoxal phosphate to lysine-214 is described; in P  
 atase activity  
 C'Comment: Serum albumin, a predominant protein in the plasma of adults, is synthesized  
 in liver, protoporphyrin, long-chain fatty acids, prostaglandins, steroid hormones (weak  
 C'Comment: A large number of variants of human serum albumin have been described.  
 C'Genetics:  
 A'Gene: GDB:ALB  
 A'Cross-references: GDB:118990; OMIM:103600  
 A'Map position: 4q11-4q13  
 C'Superfamily: serum albumin; duplication; serum albumin repeat homology  
 C'Keywords: carrier protein; signal sequence #status predicted <SIG>  
 F:1-18/Domain: signal sequence #status predicted <PRO>  
 F:19-24/Domain: propeptide #status experimental <MPT>  
 F:25-609/Product: serum albumin #status experimental <SA1>  
 F:29-202/Domain: serum albumin repeat homology <SA1>  
 F:166-174/Product: kinetensin #status experimental <KIP>  
 F:221-394/Domain: serum albumin repeat homology <SA2>  
 F:413-592/Domain: serum albumin repeat homology <SA3>  
 F:413/Binding site: copper (His) #status predicted  
 F:77-86,99-115,114-125,148-193,192-201,224-270,269-277,289-303,302-313,340-385,384-393,4  
 F:214/Binding site: pyridoxal phosphate (Lys) (covalent) #status experimental

QY 121 DVMCTAFHDNEETFLKKYLVEIARRHPYFYFAPELLFAKRYKAAFTTECCQADKAACLIP 180  
 DB 145 DVMCTAFHDNEETFLKKYLVEIARRHPYFYFAPELLFAKRYKAAFTTECCQADKAACLIP 204  
 QY 181 KLDELROBEGKASSAKQRLKASLOKFGERAFAKAWARLSORFFPKAEPFAEYSLVTLDTLK 240  
 DB 205 KLDELROBEGKASSAKQRLKASLOKFGERAFAKAWARLSORFFPKAEPFAEYSLVTLDTLK 264  
 QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 300  
 DB 265 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 324  
 QY 301 DLPSLAADFVESKOVCKNYAERAKOVFLGMFLIYEVARRHPDYSVLLILRLAKTYETLEK 360  
 DB 325 DLPSLAADFVESKOVCKNYAERAKOVFLGMFLIYEVARRHPDYSVLLILRLAKTYETLEK 384  
 QY 361 CAADAPHECYAKVDFEFKPLVEEQNLIKQNCLEFQOLGEYKFNQALILVRYTKVPOVST 420  
 DB 385 CAADAPHECYAKVDFEFKPLVEEQNLIKQNCLEFQOLGEYKFNQALILVRYTKVPOVST 444  
 QY 421 PTLVEVSRLGKVGSKCKCHPEAKRMPCAEADYLSVNLQCLVHLHEKTPVSDRVTKCCTES 480  
 DB 445 PTLVEVSRLGKVGSKCKCHPEAKRMPCAEADYLSVNLQCLVHLHEKTPVSDRVTKCCTES 504  
 QY 481 LVNRRPCPSALEVDETVYVPEKFNATFTTFHADICTLSEKERQIKQOTALVELVKKPKAT 540  
 DB 505 LVNRRPCPSALEVDETVYVPEKFNATFTTFHADICTLSEKERQIKQOTALVELVKKPKAT 564  
 QY 541 KEQLKAVNMDDFAAFVEKCKCCKADDDKTCFAEEGKKLVAASQAALGL 585  
 DB 565 KEQLKAVNMDDFAAFVEKCKCCKADDDKTCFAEEGKKLVAASQAALGL 609

RESULT 2  
 A47391  
 serum albumin precursor - rhesus macaque  
 C'Species: Macaca mulatta (rhesus macaque)  
 C'Date: 21-Jan-1994 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
 C'Accession: A47391  
 R;Watkins, S.; Sakamoto, Y.; Madison, J.; Davis, E.; Smith, D.G.; Dwulet, J.; Putnam, F.  
 Proc. Natl. Acad. Sci. U.S.A. 90, 2409-2413, 1993  
 A>Title: cDNA and protein sequence of polymorphic macaque albumins that differ in bilin  
 A'Reference number: A47391; MUID:93211971; PMID:8460152  
 A'Contents: B/B homozygote  
 A'Accession: A47391  
 A>Status: preliminary  
 A'Molecule type: mRNA; protein  
 A'Residuals: 1-600 <WAT>  
 A'Cross-references: UNIPROT:Q28522; GB:M90463; NID:G342294; PIDN:AAA36906.1; PID:G342295  
 A'Experimental source: liver  
 A>Note: sequence extracted from NCBI backbone (NCBI:128280, NCBI:128281)  
 C'Superfamily: serum albumin; serum albumin repeat homology  
 F:21-194/Domain: serum albumin repeat homology <SA1>  
 F:213-386/Domain: serum albumin repeat homology <SA2>  
 F:405-584/Domain: serum albumin repeat homology <SA3>  
 Query Match 77.7%; Score 2942; DB 2; Length 600;  
 Best Local Similarity 93.5%; Pred. No. 1.4e-69;  
 Matches 545; Conservative 23; Mismatches 15; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFKDLGEENFKALVLIIFAQYLOQCPFEDHVKLNVETFEAKTCVADESAE 60  
 DB 17 DTHKSEVAHRFKDLGEENFKGLVLVAFSQYLOQCPFEEHVKLNVETFEAKTCVADESAE 76  
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLRLVRPV 120  
 DB 77 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLRLVRPV 136  
 QY 121 DVMCTAFHDNEETFLKKYLVEIARRHPYFYFAPELLFAKRYKAAFTTECCQADKAACLIP 180  
 DB 137 DVMCTAFHDNEETFLKKYLVEIARRHPYFYFAPELLFAKRYKAAFTTECCQADKAACLIP 196

Query Match 82.0%; Score 3103; DB 1; Length 609;  
 Best Local Similarity 100.0%; Pred. No. 8.5e-74;  
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFKDLGEENFKALVLIIFAQYLOQCPFEDHVKLNVETFEAKTCVADESAE 60  
 DB 25 DAHKEVAHRFKDLGEENFKALVLIIFAQYLOQCPFEDHVKLNVETFEAKTCVADESAE 84  
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLRLVRPV 120  
 DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLRLVRPV 144



Db	204	KLDALKERILLSSAKERLKCSSFQNGFRVAKVMSVARLSQKFPKADFAEVSKIIVTDLTK	263
Qy	241	VHTECHGDLLECADRADLAKYICENODSISSSKLKECEKPLEKSHCIAEVENDEMPA	300
Db	264	VHKECHGDLLECADRADLAKYICHOQDSISGLKACCDKFLQKSHCIAEVKREDDLP	323
Qy	301	DLPSLAADPVSQKCKVNAEKVFLGMLFYEARHPDYSVLLRLAKIYETTLKCC	360
Db	324	DLPALAADPAEDKETCKYKQAKDFVLGTFLEYERRHPDYSVLLRLAKIYETTLKCC	383
Qy	361	CAAADPHCYAKVDFEFLPEEPONLJKONCELPQGLGEYKFNALLVRYTKVPQVST	420
Db	384	CAEADPPACRYTVFQDFPLVPEERKSLVKKNCDFEEVGEYDFQNALIVRYTKGAPQVST	443
Qy	421	PTLVSVSNLGVSKCKKHPRKAMPKCAEDVLSVLNQLCVLHKTVPVSRVTKCCTES	480
Db	444	PTLVIGRTLGVKGRCCCKLPESERLPCSENHALALNRLCVLHKTVPVSEKITKCTDS	503
Qy	481	LVNRRPCFSALVDETVYVPKFNPAETFTTHADICTLSKEROIKKQTMALVELVKKPKAT	540
Db	504	LAERRPCFSALVDETVYVPKFNPAETFTTHADICTLPDEKQIKKQSAALAEVLVKKPKAT	563
Qy	541	KEQLKAVMDDFRAFVEKCKKADDKETCFRAEKGKLVASQAL	583
Db	564	KEQLKTVLGNFSAFVAKCCREDKEACFAEGPKLVASSQAL	606
RESULT 5			
ABBO5			
serum albumin precursor [validated] - bovine			
N:Alternate names: 67K protein; preproalbumin			
C:Species: Bos primigenius taurus (cattle)			
C>Date: 24-Apr-1984 #sequence revision 30-Sep-1993 #text change 09-Jul-2004			
C:Accession: A38885; A36401; A91258; B60808; S10780; D45800; A26693; A90309; A91458; A94			
R:Holowachuk, E.W.; Stoltenborg, J.K.; Reed, R.G.; Peters Jr., T.			
submitted to the EMBL Data Library, August 1991			
A:Description: Bovine serum albumin: cDNA sequence and expression.			
A:Reference number: A38885			
A:Accession: A38885			
A:Molecule type: mRNA			
A:Residues: 1-607 <HOL>			
A:Cross-references: UNIPROT:P04277; EMBL:M73215			
R:Hirayama, K.; Akashi, S.; Furiya, M.; Fukuhara, K.			
Biochem. Biophys. Res. Commun. 173, 639-646, 1990			
A>Title: Rapid confirmation and revision of the primary structure of bovine serum albumi			
A:Reference number: A36401; MUID:91083649; PMID:2260575			
A:Accession: A36401			
A:Molecule type: protein			
A:Residues: 25-41,'H',43-189,'E',191-213,'T',215-323,'D',325-393,'TS',396-607 <HIR>			
R:MacGillivray, R.T.A.; Chung, D.W.; Davie, E.W.			
Eur. J. Biochem. 98, 477-485, 1979			
A>Title: Biosynthesis of bovine plasma proteins in a cell-free system.			
A:Reference number: A91258; MUID:80024278; PMID:488109			
A:Accession: A91258			
A:Molecule type: protein			
A:Residues: 1-32 <MAG>			
R:Hsieh, J.C.; Lin, F.P.; Tam, M.F.			
Anal. Biochem. 170, 1-8, 1988			
A>Title: Electrophoretic onto glass-fiber filter from an analytical isoelectrofocusing			
A:Reference number: A60808; MUID:88267456; PMID:3389500			
A:Accession: B60808			
A:Molecule type: protein			
A:Residues: 25-41 <HSI>			
R:Strawich, E.; Glincher, M.J.			
Eur. J. Biochem. 191, 47-56, 1990			
A>Title: Tooth 'enamelins' identified mainly as serum proteins. Major 'enamelin' is albu			
A:Reference number: S10780; MUID:90336641; PMID:2379503			
A:Accession: S10780			
A:Molecule type: protein			
A:Residues: 25-41,'H',43-57,59-64 <STR>			
R:Carraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.			
J. Immunol. 143, 1680-1684, 1989			
A>Title: Structures of histamine-releasing peptides formed by the action of acid protease			

A:Reference number: A45800; MUID:89341406; PMID:2474609			
A:Accession: D45800			
A:Molecule type: protein			
A:Residues: 163-172 <CAR>			
R:Carraway, R.E.; Mitra, S.P.; Cochran, D.E.			
J. Biol. Chem. 262, 5968-5973, 1987			
A>Title: Structure of a biologically active neurotensin-related peptide obtained from pe			
A:Reference number: A26693; MUID:87194805; PMID:2437111			
A:Accession: A26693			
A:Molecule type: protein			
A:Residues: 165-172,'L' <CA2>			
R:Reed, R.G.; Putnam, F.W.; Peters Jr., T.			
Biochem. J. 191, 867-868, 1980			
A>Title: Sequence of residues 400-403 of bovine serum albumin.			
A:Reference number: A90309; MUID:82023364; PMID:7283978			
A:Accession: A90309			
A:Molecule type: protein			
A:Residues: 402-433 <REE>			
R:Brown, J.R.			
Fed. Proc. 34, 591, 1975			
A>Title: Structure of bovine serum albumin.			
A:Reference number: A91458			
A:Accession: A91458			
A:Molecule type: protein			
A:Residues: 25-41,'H',43-117,'EQ',120-179,181-189,'E',191-194,'A',196-213,'T',215-288,'B'			
R:Brown, J.R.			
submitted to the Atlas, April 1975			
A:Reference number: A94551			
A:Accession: A94551			
A:Molecule type: protein			
A:Residues: 190-195 <BR2>			
R:Brown, J.R.			
Fed. Proc. 33, 1389, 1974			
A:Reference number: A91457			
A:Contents: annotation: disulfide bonds			
R:Werlen, R.C.; Offord, R.E.; Rose, K.			
Biochem. J. 302, 907-911, 1994			
A>Title: Preparation and characterization of novel substrates of insulin proteinase (EC			
A:Reference number: S55232; MUID:95031935; PMID:7945219			
A:Accession: S55232			
A>Status: preliminary			
A:Molecule type: protein			
A:Residues: 529-536;569-572 <WER>			
C:Superfamily: serum albumin; serum albumin repeat homology			
C:Keywords: carrier protein; copper binding; duplication; plasma			
F:1-18/Domain: signal sequence #status experimental <SIG>			
F:19-24/Domain: propeptide #status experimental <PRO>			
F:25-607/Product: serum albumin #status experimental <MPT>			
F:229-201/Domain: serum albumin repeat homology <SA1>			
F:220-393/Domain: serum albumin repeat homology <SA2>			
F:412-591/Domain: serum albumin repeat homology <SA3>			
F:27/Binding site: copper (His) #status predicted			
F:77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,			
Query Match 64.6%; Score 2446.9; DB 1; Length 607;			
Best Local Similarity 75.6%; Pred. No. 1.8e-56;			
Matches 441; Conservative 71; Mismatches 70; Indels 1; Gaps 1;			
Qy	1	DAKSEVAHRFKDLGEEFKALVLTAFAQYLQCCPFEDHVKLVNVEVTEFAKTCVADESAB	60
Db	25	DTKHSEIAHRFKDLGEEFKGLVLTAFAQYLQCCPFEDHVKLVNVEVTEFAKTCVADESHA	84
Qy	61	NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQPERNECFLOHKDDNPILRLVPRV	120
Db	85	GCEKSLHTLFGDELCKVASLRETYGDMADCCQKQPERNECFLSHKDDSPDLKLP	143
Qy	121	DVMCTAFHDNETFFKKLYEIAHRHPFYFAPELLFFAKRYKAAATECCQAADKAACLLP	180
Db	144	NTLCDKFADEKKFKWGLYELIARRHPFYFAPELLFYANKYNGVFQDCCQAEKGCALLP	203
Qy	181	KLDELDEGKASSAKORLCKASLOKFGRAFKAWAVARLSORFFKAEFAEVSKLVTDLTK	240
Db	204	KIETWREKVLASSARQRURCASIQFGKALKAWSVARLSQKFAEFVEVTKLVTDLTK	263

```
Qy 241 VHTCCHGDLLECCADRADLAKYICENODSISXKLECCCKPILKESHCIAEVENDEMPA 300
Db 264 VHKECCHGDLLECCADRADLAKYICDNQDTISSKLECCCDKPLLEKSHCIAEVEKDAIPE 323
Qy 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFYIYARRHPDYSVVLLRLAKTYETTTLEKC 360
Db 324 NLPLTADPAEDKDVCKNYQEAQDAFLGSLFYIYARRHPDYSVVLLRLAKTYETTTLEKC 383
Qy 361 CAADPHCEYAKVDFEFKPLVEBPONLIKONCELFEOLGEYKFNQALLVRYTKVPQVST 420
Db 384 CAKDDPHACVATVDFDKHLVDEPQNLIKONCELFEKLGGEYGFQNALIVRYTKVPQVST 443
Qy 421 PTLVEVSRNLGVSKCKCHPEAKRMPCAEDYLSVNLQCLVLEHKTTPVSDRVTKCCTES 480
Db 444 PTLVEVSRNLGVSKCKCHPEAKRMPCTEDYLSLILNRLCVLHKTTPVSEKVTKCCTES 503
Qy 481 LVNRRPCFSALVEDETVVPKFNATFTTFHADICTLSEKEROIKQKOTALVELVVKHKPKAT 540
Db 504 LVNRRPCFSALTPTDETVVPKAFDEKLTFTFHADICTLPTDEKQIKQOTALVELVVKHKPKAT 563
Qy 541 KEOLKAVMDPFAAFVEKCKKADDETCFAEGGKLVAAASQAAL 583
Db 564 EQLKVTWENFVAFVDRKCAADDEKFAVEGPKLVVSTQAL 606

RESULT 6
ABSHS
serum albumin precursor - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C:Accession: S06936
R:Brown, W.M.; Dziegielewska, K.M.; Foreman, R.C.; Saunders, N.R.
Nucleic Acids Res. 17, 10495, 1989
A:Title: Nucleotide and deduced amino acid sequence of sheep serum albumin.
A:Reference number: S06936; UID:9009888; PMID:2602160
A:Accession: S06936
A:Molecule type: mRNA
A:Residues: 1-607 <BRO>
A:Cross-references: UNIPROT:P14639; EMBL:X17055; NID:G1386; PIDN:CAA34903.1; PID:G1387
C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,
teroid hormones (weak bonds with these hormones promote their transfer across the membra
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: carrier protein; duplication; metal binding; plasma
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-24/Domain: propeptide #status predicted <PRO>
F:25-607/Product: serum albumin #status predicted <MAT>
F:29-201/Domain: serum albumin repeat homology <SA1>
F:220-393/Domain: serum albumin repeat homology <SA2>
F:412-591/Domain: serum albumin repeat homology <SA3>
F:27/Binding site: copper (His) #status predicted
F:77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,4
F:263/Binding site: bilirubin (Lys) #status predicted

Query Match 64.3%; Score 2432.9; DB 1; Length 607;
Best Local Similarity 75.0%; Pred. No. 4.1e-56;
Matches 437; Conservative 73; Mismatches 72; Indels 1; Gaps 1;

Qy 1 DAHSEVAHRFKDLGEENFKALVIAFAQYLQCCPFEDHVKLVNVEVTEFAKTCVADSESAE 60
Db 25 DTHKSEIAHRNDLGEENFQGLVIAFSQYLQCCPFDEHVKLVKELTEFAKTCVADSESHA 84
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPFLRLVRPEV 120
Db 85 GCDKSLHTLFGDELCKVATLRETYGDMADCCQKQEPERNECFNLNKHKDSPLKLRPEP 143
Qy 121 DVMCTAFHDNETFLKYLVTARHPYFYAPELLFFAKRYKAAFTCCQAAADKAACILP 180
Db 144 DTLCAEFKADKKFKWGKLYIEVARHPYFYAPELLTYANKNGVFOECQADKGAACILP 203
Qy 181 KLDELDRDGKASSAKQRLKCSLQKFGERFAKAVARLSQRFPAEFAVSKLVTDLTK 240
Db 204 KIDAMREKVLASSARQRLCASIQKFGERALKAWSVARLSQKFPKADFTDVTKIIVTDLTK 263
```

```
Qy 241 VHTCCHGDLLECCADRADLAKYICENODSISXKLECCCKPILKESHCIAEVENDEMPA 300
Db 264 VHKECCHGDLLECCADRADLAKYICDQDASXKLECCCKPILKESHCIAEVDKDAVPE 323
Qy 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFYIYARRHPDYSVVLLRLAKTYETTTLEKC 360
Db 324 NLPLTADPAEDKDVCKNYQEAQDAFLGSLFYIYARRHPDYSVVLLRLAKTYETTTLEDC 383
Qy 361 CAADPHCEYAKVDFEFKPLVEBPONLIKONCELFEOLGEYKFNQALLVRYTKVPQVST 420
Db 384 CAKDDPHACVATVDFDKHLVDEPQNLIKONCELFEKLGGEYGFQNALIVRYTKAPQVST 443
Qy 421 PTLVEVSRNLGVSKCKCHPEAKRMPCAEDYLSVNLQCLVLEHKTTPVSDRVTKCCTES 480
Db 444 PTLVEVSRNLGVSKCKCHPEAKRMPCTEDYLSLILNRLCVLHKTTPVSEKVTKCCTES 503
Qy 481 LVNRRPCFSALVEDETVVPKFNATFTTFHADICTLSEKEROIKQKOTALVELVVKHKPKAT 540
Db 504 LVNRRPCFSALTPTDETVVPKAFDEKLTFTFHADICTLPTDEKQIKQOTALVELVVKHKPKAT 563
Qy 541 KEOLKAVMDPFAAFVEKCKKADDETCFAEGGKLVAAASQAAL 583
Db 564 EQLKVTWENFVAFVDRKCAADDEKFAVEGPKLVVSTQAL 606

RESULT 7
ABRTS
serum albumin precursor - rat
N:Alternate names: preproalbumin
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-May-1979 #sequence_revision 31-May-1979 #text_change 09-Jul-2004
C:Accession: A93872; A92211; A91946; A91940; C45800; I57621; A03233
R:Sargent, T.D.; Yang, M.; Bonner, J.
Proc. Natl. Acad. Sci. U.S.A. 78, 243-246, 1981
A:Title: Nucleotide sequence of cloned rat serum albumin messenger RNA.
A:Reference number: A93872; UID:8122372; PMID:7017712
A:Accession: A93872
A:Molecule type: mRNA
A:Residues: 1-608 <SAR>
A:Cross-references: UNIPROT:P02770; GB:V01222; GB:J00698; NID:955627; PIDN:CAA24532.1; PPI:
R:Strauss, A.W.; Bennett, C.D.; Donohue, A.M.; Rodkey, J.A.; Alberts, A.W.
J. Biol. Chem. 252, 6846-6855, 1977
A:Title: Rat liver pre-proalbumin: complete amino acid sequence of the pre-piece. Analysis
A:Reference number: A92211; UID:77249657; PMID:893447
A:Note: cleavages during protein maturation
A:Accession: A92211
A:Molecule type: protein
A:Residues: 1-38 <STR>
R:Isemura, S.; Ikenaka, T.
J. Biochem. 83, 35-48, 1978
A:Title: Amino acid sequences of fragments I and II obtained by cyanogen bromide cleavage
A:Reference number: A91946; UID:78109429; PMID:564345
A:Accession: A91946
A:Molecule type: protein
A:Residues: 25-222 <ISI>
R:Isemura, S.; Ikenaka, T.
J. Biochem. 79, 1183-1196, 1976
A:Title: Fragmentation of rat serum albumin by cyanogen bromide cleavage and the amino a
A:Reference number: A91940; UID:76260153; PMID:956149
A:Accession: A91940
A:Molecule type: protein
A:Residues: 223-288; 572-608 <IS2>
A:Note: 262-Leu was also found
R:Aoyagi, Y.; Ikenaka, T.; Ichida, F.
Cancer Res. 38, 3483-3486, 1978
A:Title: Copper(II)-binding ability of human alpha-fetoprotein.
A:Reference number: A90758; UID:79001617; PMID:802665
A:Contents: annotation; copper binding
R:Carraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.
J. Immunol. 143, 1680-1684, 1989
A:Title: Structures of histamine-releasing peptides formed by the action of acid protease
A:Reference number: A45800; UID:89341406; PMID:2474609
```

A:Accession: C45800  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 166-173 <AR>  
R:Heard, J.

Mol. Cell. Biol. 7, 2425-2434, 1987  
A>Title: Determinants of rat albumin promoter tissue specificity analyzed by an improved  
A:Reference number: 157621; MUID:87286876; PMID:3475566

A:Accession: I57621

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-5 <RES>

A:Cross-references: GB:M16825; NID:g202828; PIDN:AAA0712.1; PID:g554412

C:Superfamily: serum albumin; serum albumin repeat homology

C:Keywords: carrier protein; duplication; metal binding; plasma

F:1-18/Domain: signal sequence #status experimental <SIG>

F:19-24/Domain: propeptide #status experimental <PRO>

F:25-608/Product: serum albumin #status experimental <MAT>

F:29-202/Domain: serum albumin repeat homology <SA1>

F:221-394/Domain: serum albumin repeat homology <SA2>

F:413-592/Domain: serum albumin repeat homology <SA3>

F:27/Binding site: copper (His) #status experimental

F:77-86, 99-115, 114-125, 148-193, 192-201, 224-270, 269-277, 289-303, 302-313, 340-385, 384-393, 4

Query Match 64.1%; Score 2426; DB 1; Length 608;

Best Local Similarity 73.4%; Pred. No. 6.3e-56;

Matches 428; Conservative 82; Mismatches 73; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGSENFKALVLIAPAYLOQCPEFHVKLVNEVTEPAKTCVADESAB 60

DB 25 EAHKSEIAHRFKDLGSENFKALVLIAPAYLOQCPEFHVKLVNEVTEPAKTCVADESAB 84

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQPERNECFLOHKDNDNPNLPRVVRPV 120

DB 85 NCDKSIHTLFGDKLCAIPSLREHYGDLADCCBKEPERNECFLOHKDNDNPNLPPQRPBA 144

QY 121 DVMTAFHDNEFTFLKKLYELIARRHPYFYAPPELLFPAKRYKAAFTCCQAAADKAACLTP 180

DB 145 EAMCTSFQENPTSFGLHYLHEVARHPYFYAPPELLYAEKYNEVLVTQCTESDKAACLTP 204

QY 181 KLDELDRDEGKASSAKORLKCSAQKGFERAFKAWAVARLSQRPFAEVSKLVTDLTK 240

DB 205 KLDVAKKEKALVAARQMKCSSQRFGERAFKAWAVARMSQRPFAEVSKLVTDLTK 264

QY 241 VHTCECHGDLLECCADRADLAKYICENODSISKLKECEKPELLEKSHCHIAEVDNDMPA 300

DB 265 INKECHGDLLECCADRADLAKYICENQATISKLQACCDKPVQLQSQCLAEVHDNTPA 324

QY 301 DLPSLAADPFVESKDVCKNYAEAKDVFGLMFLYEYARRHPDYSVLLRLAKTYETTELEK 360

DB 325 DLPSIAADPFVEDKEVCKNYAEAKDVFGLTFLYEYSRRHPDYSVLLRLAKTYETTELEK 384

QY 361 CAAADPHECYAKVDFDFKPLVEEPQNLIKQNCFLFQGEYKFNALLVRYTKVPQVST 420

DB 385 CAEGDPPACYATVDFKQPLVDEPKNLIKQNCFLFQGEYKFNALLVRYTKVPQVST 444

QY 421 PTLVEVRNLGVKSGKCKKHPAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480

DB 445 PTLVEARNLGRVGTAKCTLPPEAQRUPCEDYLSAIIANLRLCVLHEKTPVSEKTKCCSGS 504

QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVGHKPKAT 540

DB 505 LVRRPCFSALTVDETYVPKEFNAETFTFHSDICTLPDKEKQIKKOTALVELVGHKPKAT 564

QY 541 KQOLKAVMDDFAAVFEKCKKADDKETCFABEGKLVAAQAAAL 583

DB 565 EDQLTKVMGDFAQFVDKCKCAADKDNCFATEGPNLVARSKAAL 607

RESULT 8

ABPGS

serum albumin precursor - pig (fragment)

C:Species: Sus scrofa domestica (domestic pig)

C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004

C:Accession: S01382; A61006

R:Weinstock, J.; Baldwin, G.S.

Nucleic Acids Res. 16, 9045, 1988

A>Title: Nucleotide sequence of porcine liver albumin.

A:Reference number: S01382; MUID:89016582; PMID:3174440

A:Accession: S01382

A>Status: translation not shown

A:Molecule type: mRNA

A:Residues: 1-605 <WEI>

A:Cross-references: UNIPROT:P08835; EMBL:X12422; NID:g1875; PIDN:CAA30970.1; PID:g833798

R:Limeback, H.; Sakarya, H.; Chu, W.; Mackinnon, M.

J. Bone Miner. Res. 4, 235-241, 1989

A>Title: Serum albumin and its acid hydrolysis peptides dominate preparations of mineral

A:Reference number: A61006; MUID:89269769; PMID:2728927

A:Accession: A61006

A:Molecule type: protein

A:Residues: 23-51, 'X', 53-54, 'XXXGY', 146, 'E', 148, 'E', 150-151, 'XVN', 155 <LIM>

A:Experimental source: dental enamel

A>Note: albumin and other serum proteins are also found in bone

C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,

teroid hormones (weak bonds with these hormones promote their transfer across the membra

C:Superfamily: serum albumin; serum albumin repeat homology

C:Keywords: carrier protein; duplication; metal binding; plasma

F:1-16/Domain: signal sequence (fragment) #status predicted <SIG>

F:17-22/Domain: propeptide #status predicted <PRO>

F:23-605/Product: serum albumin #status predicted <MAT>

F:27-199/Domain: serum albumin repeat homology <SA1>

F:218-391/Domain: serum albumin repeat homology <SA2>

F:410-589/Domain: serum albumin repeat homology <SA3>

F:75-84, 97-113, 112-123, 145-190, 189-198, 221-267, 266-274, 286-300, 299-310, 337-382, 381-390,

F:261/Binding site: bilirubin (Lys) #status predicted

Query Match 63.9%; Score 2416.8; DB 1; Length 605;

Best Local Similarity 71.8%; Pred. No. 1.1e-55;

Matches 441; Conservative 70; Mismatches 71; Indels 33; Gaps 2;

QY 1 DAHSEVAHRFKDLGSENFKALVLIAPAYLOQCPEFHVKLVNEVTEPAKTCVADESAB 60

DB 23 DTYKSEIAHRFKDLGSENFKALVLIAPAYLOQCPEFHVKLVNEVTEPAKTCVADESAB 82

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQPERNECFLOHKDNDNPNLPRVVRPV 120

DB 83 NCDKSIHTLFGDKLCAIPSLREHYGDLADCCBKEPERNECFLOHKDNDNPNLPRVVRPV 141

QY 121 DVMTAFHDNEFTFLKKLYELIARRHPYFYAPPELLFPAKRYKAAFTCCQAAADKAACLTP 180

DB 142 VALCADFOEDGQKFWGKLYELIARRHPYFYAPPELLYALIIYKDFSECCQAAADKAACLTP 201

QY 181 KLDELDRDEGKASSAKORLKCSAQKGFERAFKAWAVARLSQRPFAEVSKLVTDLTK 240

DB 202 KIEHLREKVLTSAAQRLKCSAQKGFERAFKAWAVARLSQRPFAEVSKLVTDLTK 261

QY 241 VHTCECHGDLLECCADRADLAKYICENODSISKLKECEKPELLEKSHCHIAEVDNDMPA 300

DB 262 VHTCECHGDLLECCADRADLAKYICENQDTISTKLKECCDKPPELLEKSHCHIAEVDNDMPA 321

QY 301 DLPSLAADPFVESKDVCKNYAEAKDVFGLMFLYEYARRHPDYSVLLRLAKTYETTELEK 360

DB 322 DLNPLEHDFVEDKEVCKNYAEAKDVFGLTFLYEYSRRHPDYSVLLRLAKTYETTELEK 381

QY 361 CAAADPHECYAKVDFDFKPLVEEPQNLIKQNCFLFQGEYKFNALLVRYTKVPQVST 420

DB 382 CAKEDPPACYATVDFKQPLVDEPKNLIKQNCFLFQGEYKFNALLVRYTKVPQVST 441

QY 421 PTLVEVRNLGVKSGKCKKHPAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480

DB 442 PTLVEVARNLGLVSGRCKRPEEELSCAEDYLSVVLNQLCVLHEKTPVSEKTKCTES 501

QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVGHKPKAT 540

DB 502 LVNRRPCFSALTVDETYVPKEFVEGFTTFHADICTLPDEKQIKKOTALVELVGHKPKAT 561

```
Qy 541 KEQLKAVMDPFAAFVEKCKADDKETCFABEGKKLVAASQAALGLMSPRLEVPCHALPQ 600
Db 562 EEQLRTVLGNFAAFVQKCAAPDHEACFAVEGPKF-----596
Qy 601 GLSPGQVIVRGLV 614
Db 597 -----VIEIRGIL 604

RESULT 9
JC5838
serum albumin - Mongolian jird
C;Species: Meriones unguiculatus (Mongolian jird)
C;Date: 05-Mar-1998 #sequence_revision 13-Mar-1998 #text_change 09-Jul-2004
C;Accession: JC5838
R;Ioshida, K.; Sero-Oshima, A.; Sinohara, H.
DNA Res. 4, 351-354, 1997
A;Title: Sequencing of cDNA encoding serum albumin and its extrahepatic synthesis in the
A;Reference number: JC5838; MUID:98116663; PMID:9455485
A;Molecule type: mRNA
A;Residues: 1-609 <YOS>
A;Cross-references: UNIPROT:O35090; DBJ:AB006197; NID:g2317277; PIDN:BAA21765.1; PID:g2317277
A;Experimental source: liver
C;Superfamily: serum albumin; serum albumin repeat homology
F;222-395/Domain: serum albumin repeat homology <SA2>

Query Match 63.1%; Score 2387; DB 2; Length 609;
Best Local Similarity 73.9%; Pred. No. 6.8e-55;
Matches 430; Conservative 65; Mismatches 87; Indels 0; Gaps 0;

Qy 2 AHKSEVARRFKDLGEENFKALVIAFAQYLQOCPEFDHVKLVNVEVTRFAKTCVADSAEN 61
Db 27 AHKSEIAHRYKDLGEKFKGLVLTFFSQYLQKCYEEHVKLVREVTDFASNCADSAEN 86
Qy 62 CDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPMLPRLVLRPEVD 121
Db 87 CDKSLHTLFGDKLCLNPFGEKYAEMADCCAKQEPERNECFLOHKDNPFLPPFKRAEPD 146
Qy 122 VMCTAFHDNETFKKYLVEIARRHPYFYAPPELLFFAKRYKAAATECCQAADKAACLLPK 181
Db 147 AMCTAFQENAEAFMGHYLHEVARRHPYFYGPPELLYLDKVTAVLTTECCAADDKGACLTTPK 206
Qy 182 LDELDEGKASSAKORLKCSAQKQEGRAFKAWAVARLSORFPKAEAEVSKLVTDLTKV 241
Db 207 LDALKEKALVSARQRLKCSSMKKFGRAFKAWAVARMSQTFPNADFAEITKGLATDTKV 266
Qy 242 HTECHGDLLECADRDADLAKYICENQDSISSKLKECEKEPILLEKSHCIAEVENDEMPAD 301
Db 267 TQECHGDLLECADRDADLAKYKMCENQASISSKLQACCDKEMLKQSCQLAEVEHDDMPAD 326
Qy 302 LPSLAADPVESKDYCKNYAEAKDVLGMFLVEYARRHPDYSVVLRLRLAKTYETTLBKCC 361
Db 327 LPALTADFVEDKDYCKNYAEAKDVLGTFLYEYSRRHPDYSVLSLLRLAKKYEATLEKCC 386
Qy 362 AAADPHCYAKVDFEFPKLVPEEPONLTKONCELFQELGEYKFNALLVRYTKKVPQVSTP 421
Db 387 AEADPHACYGHVDFEFPKLVPEEPONLVKSNCELYEKLGEYFQNAVLRVYTKKAPQVSTP 446
Qy 422 TLVEVSNRLKGVGSKCKHPKAKMPCAEEDYLSVVLNOLCVLHEKTPVSDRVTKCCTESL 481
Db 447 TLVFAARSLGRVTHCCALPEKKRLPCVEDYLSAILNREVCLLHKEKTPVSEQVTKCCSGSL 506
Qy 482 VNRRPCFSALVEVDYTPKFNFAETFTFHADICTLSEKERQIKKQATLVELVHKPKATK 541
Db 507 VERRPCFSALPVDEYTPKFKFAETFTFHANICTLPEKEKQMEKQATALAEVLVHKHPQATE 566
Qy 542 EQLKAVMDPFAAFVEKCKADDKETCFABEGKKLVAASQAAL 583
Db 567 EQLKAVMDPFAAFVEKCKQEDKEACFSTEGPKLVAESQKAL 608

RESULT 10
```

```
A05139
serum albumin - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 05-Jun-1987 #sequence_revision 17-Mar-2000 #text_change 09-Jul-2004
C;Accession: A05139; I48638
R;Minghetti, P.P.; Law, S.W.; Dugaiczky, A.
Mol. Biol. Evol. 2, 347-358, 1985
A;Title: The rate of molecular evolution of alpha-fetoprotein approaches that of pseudog
A;Reference number: A93055; MUID:88216123; PMID:2452956
A;Accession: A05139
A;Molecule type: mRNA
A;Residues: 1-418 <MIN>
A;Cross-references: UNIPROT:P07724; GB:M16111; NID:g191764; PIDN:AAA37190.1; PID:g191765
R;Boccaccio, C.; Deschatrette, J.; Meunier-Rotival, M.
Gene 88, 181-186, 1990
A;Title: Empty and occupied insertion site of the truncated LINE-1 repeat located in the
A;Reference number: I48638; MUID:90289606; PMID:1971802
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 379-453 <BOC>
A;Cross-references: EMBL:X13060; NID:g52939; PIDN:CAA31458.1; PID:g899334
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: carrier protein; duplication; metal binding; plasma
F;1-104/Domain: serum albumin repeat homology (fragment) <SA1>
F;123-296/Domain: serum albumin repeat homology <SA2>
F;315-453/Domain: serum albumin repeat homology (fragment) <SA3>

Query Match 49.2%; Score 1861; DB 2; Length 453;
Best Local Similarity 72.2%; Pred. No. 2e-41;
Matches 327; Conservative 64; Mismatches 62; Indels 0; Gaps 0;

Qy 75 CTVATLRETYGEMADCCAKQEPERNECFLOHKDNPMLPRLVLRPEVDVMCTAFHDNETF 134
Db 1 CAIPNLRKENTGELADCCCKQEPERNECFLOHKDNPMLPRLVLRPEVDVMCTAFHDNETF 60
Qy 135 LKLYEYIARRHPYFYAPPELLFFAKRYKAAATECCQAADKAACLLPKLDELDEGKASSA 194
Db 61 MGHYLHEVARRHPYFYAPPELLYLYAEQYNEILTQCCAEDKESCLTPKLDGVKEKALVSSV 120
Qy 195 KQRLKASLQKQGERAFKAWAVARLSORFPKAEAEVSKLVTDLTKVHTCCCHGDLLECA 254
Db 121 RQRMKCSSMKQKGERAFKAWAVARLSQTFPNADFAEITKGLATDTLVNKECCHGDLLECA 180
Qy 255 DDRADLAKYICENQDSISSKLKECEKEPILLEKSHCIAEVENDEMPADLPSLAADPVESKD 314
Db 181 DDAELAKYKMCENQATISSKLQTCDDKPLKKAHCLSEVEHDTMPADLPAIADFVEDQE 240
Qy 315 VCKNYAEAKDVLGMFLVEYARRHPDYSVVLRLRLAKTYETTLKCCAAADPHCYAKVF 374
Db 241 VCKNYAEAKDVLGTFLYEYSRRHPDYSVLSLLRLAKKYEATLEKCCAEANPPACYGTVL 300
Qy 375 DEFKPLVEEPONLTKONCELFQELGEYKFNALLVRYTKKVPQVSTPTLVEVSNRLKGV 434
Db 301 AEFQPLVEEPKLVKTNCDLYEKLGEYGFQNAVILRYTKAPQVSTPTLVEAEANLGRVG 360
Qy 435 SKCCHEPEAKMPCAEEDYLSVVLNOLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALVEVD 494
Db 361 TKCTLPEDQRLPCVEDYLSAILNREVCLLHKEKTPVSEHVTKCCSGSLVRRPCFSALTVD 420
Qy 495 ETVYVPKFNFAETFTFHADICTLSEKERQIKKQT 527
Db 421 ETVYVPKFNFAETFTFHSDICTLPEKEKQIKKQT 453

RESULT 11
ABCS8
serum albumin precursor - chicken
C;Species: Gallus gallus (chicken)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: S15571; A05078; A13451
R;Cassady, A.I.; Salkild, C.K.; Baverstock, P.; Wallace, J.C.
submitted to the EMBL Data Library, July 1991
```



A;Reference number: S15571  
A;Accession: S15571  
A;Molecule type: mRNA  
A;Residues: 1-615 <CAS>  
A;Cross-references: UNIPROT:P19121; EMBL:X60688; NID:G63747; PIDN:CAA43098.1; PID:G63748  
R;Hache, R.J.G.; Wiskocil, R.; Vasa, M.; Roy, R.N.; Lau, P.C.K.; Deeley, R.G.  
J. Biol. Chem. 258, 4556-4564, 1983  
A;Title: The 5' noncoding and flanking regions of the avian very low density apolipoprotein  
A;Reference number: A05078; MUID:83161037; PMID:6187737  
A;Accession: A05078  
A;Molecule type: DNA  
A;Residues: 1-28 <HAC>  
A;Cross-references: GB:V00381; NID:G63038; PIDN:CAA23680.1; PID:G63039  
R;Rosen, A.M.; Geller, D.M.  
Biochem. Biophys. Res. Commun. 78, 1060-1066, 1977  
A;Title: Chicken microsomal albumin: amino terminal sequence of chicken proalbumin.  
A;Reference number: A13451; MUID:78019943; PMID:911327  
A;Accession: A13451  
A;Molecule type: protein  
A;Residues: 19-23, 'M', 25-30 <ROS>  
C;Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper, mones (weak bonds with these hormones promote their transfer across the membranes), thyronines  
C;Superfamily: serum albumin; serum albumin repeat homology  
C;Keywords: carrier protein; duplication; metal binding; plasma  
F;1-18/Domain: signal sequence #status predicted <SIG>  
F;19-26/Domain: propeptide #status predicted <PRO>  
F;27-613/Product: serum albumin #status predicted <MAT>  
F;32-206/Domain: serum albumin repeat homology <SA1>  
F;32-206/Domain: serum albumin repeat homology <SA2>  
F;417-596/Domain: serum albumin repeat homology <SA3>  
F;30/Binding site: copper (His) #status predicted  
F;80-89, 102-118, 117-128, 152-197, 196-205, 228-274, 273-281, 293-307, 306-317, 344-389, 388-397,  
Query Match 41.2%; Score 1557.9; DB 1; Length 615;  
Best Local Similarity 46.7%; Pred. No. 5.4e-33;  
Matches 273; Conservative 118; Mismatches 192; Indels 1; Gaps 1;  
QY 3 HKSEVAHRFKDLEENFKALVLIAPAYQLQCCPFEDHVKLVNNEVTEFAKTQVADSAENC 62  
DB 30 HKSEIAHRYNDLKEETFRVAMITTPAYLQRCSEGLSKLVKVDVLAQKCVANEDAPEC 89  
QY 63 KSLHTLFGDKLCTVATLRETYGENADCCAKQEPERNECFLOHKDDNP-LVRPEVD 121  
DB 90 SKPLPSIILDETCQVEKLRDSYGAMADCCSKADPERNECFLSFKYSQDPFVQYQPSAD 149  
QY 122 VMCTAFDNEETFLKYLVEIARRHPYFVAPPELLFAKRYKAAFTCCQAAKAACLLPK 181  
DB 150 VICQYQDNRVSLGHFIYSVARRHPFLYAPILLSFVDFHALQSCCKESDVGACLDTK 209  
QY 182 LDELDEGKASSAKQRLKCSAQKFGERAFKAWAVARLSQRPFAEVSXKLVLDLTKV 241  
DB 210 EIVMEKAGVSKQYFCGILKQFGDRVQFARQLIYLSQYKPAFSEVSKFVHDSIGV 269  
QY 242 HTECCGDLLECCADRADLAKYICNQDSISSKLEKCECKPLLEKSHGICIAEVNDEMAD 301  
DB 270 HKECEGDMVECDMDMARMSNLCSQDDVFSKIKDCCEKPIVERSCQIMEAEFDEKPAD 329  
QY 302 LPSLAADFVSKDVCKNYAEAKDVLGMLFVEYARRHPDYSVVLLRLAKTETTLKCC 361  
DB 330 LPSLVEKYIEDKVECKSFAGHDAFMAEFVYSRRHPFSQLIMRIAKGYESLLEKCC 389  
QY 362 AAADPHECVAKYFDEFKPLVEBPQNLIKONCELFQGLGEYKFNQALLVRYTKKVPQVSTP 421  
DB 390 KTDNPAECYANAQEQNLQHIKETQDVVKTNCDDLLHDHGEADFLKSLIRYTKKMQVPTD 449  
QY 422 TLVVEVRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTESL 481  
DB 450 LLETKGKMTTIGTKCCQLGDRRRMACSEGYLSIVIHDTCRKQETTPINDNVSQCCSQLY 509  
QY 482 VNRRPCFSALEVDYTVPKFNAETFTHADICTLSEKERQIKKQTALVELVKKHPKATK 541  
DB 510 ANRRPCFTAMGVDTKYVPPFPFDMFSDFKLCSAPAEERGVQMKLLINLIKRPQMTPE 569

QY 542 EQLKAVMDPFAAFVEKCKCKADDKETCFABEGKULVAASQAALGL 585  
DB 570 EQIKTIADGFTAMVDKCKCKQSDINTCFGEGANLIVQSRATLGI 613  
RESULT 12  
JC4258  
alpha-fetoprotein precursor - chimpanzee  
C;Species: Pan troglodytes (chimpanzee)  
C;Date: 27-Nov-1995 #sequence\_revision 08-Feb-1996 #text\_change 09-Jul-2004  
C;Accession: JC4258  
R;Nishio, H.; Gibbs, P.E.M.; Minghetti, P.P.; Zielinski, R.; Dugaiczky, A.  
Gene 162, 213-220, 1995  
A;Title: The chimpanzee alpha-fetoprotein-encoding gene shows structural similarity to  
A;Reference number: JC4258; MUID:96032345; PMID:7557431  
A;Accession: JC4258  
A;Molecule type: DNA  
A;Residues: 1-609 <NIS>  
A;Cross-references: UNIPROT:Q28789; GB:U21916; NID:G841311; PIDN:AAA91641.1; PID:G841312  
C;Comment: This protein is a plasma protein produced in the fetal and neonatal liver an  
o similar properties and structure.  
C;Genetics:  
A;Gene: afp  
A;Map position: 3p  
A;Introns: 29/1; 46/2; 90/3; 161/2; 205/3; 238/2; 281/3; 353/2; 397/3; 430/2; 476/3; 551  
C;Superfamily: serum albumin; serum albumin repeat homology  
C;Keywords: glycoprotein  
F;1-19/Domain: signal sequence #status predicted <SIG>  
F;20-609/Product: alpha-fetoprotein #status predicted <MAT>  
F;29-202/Domain: serum albumin repeat homology <SA1>  
F;221-394/Domain: serum albumin repeat homology <SA2>  
F;413-592/Domain: serum albumin repeat homology <SA3>  
F;42,251/Binding site: carbohydrate (Asn) #status predicted  
Query Match 33.2%; Score 1256.7; DB 2; Length 609;  
Best Local Similarity 39.9%; Pred. No. 4.7e-25;  
Matches 237; Conservative 113; Mismatches 231; Indels 13; Gaps 3;  
QY 4 KSEVAHR-----FKDLGEENFKALVLIAPAYQLQCCPFEDHVKLVNNEVTEFAKT 52  
DB 17 ESRTLHREYGIASLDSYQCTAEINLTDLATIPFAQVQEAITYKEVSQWMDALTAISK 76  
QY 53 CVADESANCKSLHTLFGDKLCTVATLRETYGENADCCAKQEPERNECFLOHKDDNP-N 111  
DB 77 PTGDRSQAGCLENQLPAFLEELCREKILEKYGH-SDCCSQSEGEHRNCFLAHKPTPAS 135  
QY 112 LPLRVPEVDVMCTAFDNEETFLKYLVEIARRHPYFVAPPELLFAKRYKAAFTCCCOA 171  
DB 136 IPFFQVPEPVTSCAEYEDRETFFMKNFIYETARRHPFLYAPILLWAARYDKLIIPSCCKA 195  
QY 172 ADKAAACLLPKLDELDEGKASSAKQRLKCSAQKFGERAFKAWAVARLSQRPFAEFAEV 231  
DB 196 ENAVECFQTKAATVTKELRESSLLNQHACAVMKNFGTRTFQAITVTKLSQKFTKVNFTEI 255  
QY 232 SKLVTDLTKVHTECCGDLLECCADRADLAKYICNQDSISSKLEKCECKPLLEKSHGICIA 291  
DB 256 QKLVLDVAHVHEHCRCGDVLDCLQDGEKIMSYICSQDDTSLNKITCECKLTLERGCQII 315  
QY 292 EVNDEMADLPSLAADFVSKDVCKNYAEAKDVLGMLFVEYARRHPDYSVVLLRLAK 351  
DB 316 HAENDEKPEGLSPNLRFLGDRDNQFSGEKNIFLASFVHYSRRHPQLAVSVILRVAK 375  
QY 352 TYETTLKCCAAADPHECVAKYFDEFKPLVEBPQNLIKONCELFQGLGEYKFNQALLVRY 411  
DB 376 GYOELLEKCFQTEENPLEQDKGEEELQKYIQESQALAKRSCGLFQKLGYYLQNAFLVAY 435  
QY 412 TKKVPQVSTPTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSD 471  
DB 436 TKKAPQLTSSELMAITRMAATAATCCQLSEDKLLACGEAGADIIGHLCIRHETTPVNP 495  
QY 472 RYTKCTESLVNRRPCFSALEVDYTVPKFNAETFTHADICTLSEKERQIKKQTALVELV 531  
DB 496 GVGQCCTSSYANRRPCFSSLVVDYTVPPAFSDDKFI FHKDLQCSAQGVALQTMKQEFLLIN 555





```

QY 352 TYETTLKCCAAADPHCYAKVDFEKLVPBEPQNLIKONCELPBQLGEYKFNALLVRY 411
   : ||||| : ||| : : : : : ||||| : ||||| : ||||| : ||||| : |||||
Db 376 GYQELLEKCFQTNPLECQDKGEBELQKYIQESQALAKRSCGLFQKLGUYLQNAFLVAY 435
   : ||||| : ||| : : : : : ||||| : ||||| : ||||| : ||||| : |||||
QY 412 TKKYPQVSTPTLTVESVNLGVGSKCKHPKAPKMPCAEDYLSVVLNOLCVLHKTPTVSD 471
   ||| ||||| : : : : : ||||| : ||||| : ||||| : ||||| : |||||
Db 436 TKKAPQLTSSSELMATRKMAATAATCCQLSEDKLLACGEGAADIIGHLCIRHEMTVPNP 495
   ||| ||||| : : : : : ||||| : ||||| : ||||| : ||||| : |||||
QY 472 RVTKCTESLVNRRPCFSALEVDETVYPKFNAAETFTPHADICTLSEKERQIKKQTALVE 531
   : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
Db 496 GVGQCTSSYANRRPCFSLVVDYTVPPAFSDDKFIFHKOLCOAQGVALTQMKQEFLLN 555
   : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
QY 532 LVKHKPKATKQLKAVMDPFAFVEKCKCKADDKTCFAEBGKLVAAQAALGL 585
   ||| ||||| : : : : : ||||| : ||||| : ||||| : ||||| : |||||
Db 556 LVKQKPQITEGLEAVIADFGLLEKCCQGQEQVEVCFABEGQKLSKTRAAALGV 609

RESULT 14
FFGO
alpha-fetoprotein precursor - gorilla
C:Species: Gorilla gorilla (gorilla)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C:Accession: A37970
R:Ryan, S.C.; Zielinski, R.; Dugaiczky, A.
Genomics 9, 60-72, 1991
A:Title: Structure of the gorilla alpha-fetoprotein gene and the divergence of primates.
A:Reference number: A37970; MUID:91169517; PMID:1706310
A:Accession: A37970
A:Molecule type: DNA
A:Residues: 1-609 <RYA>
A:Cross-references: UNIPROT:P28050; GB:M38272; NID:9817963; PIDN:AAA73520.1; PID:gl77041
C:Genetics:
A:Map position: 4q11-12
A:Introns: 29/1; 46/2; 205/3; 161/2; 281/3; 353/2; 397/3; 430/2; 476/3; 551
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: embryo; fetus; globulin; glycoprotein; metal binding; plasma
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-609/Product: alpha-fetoprotein #status predicted <MAT>
F:29-202/Domain: serum albumin repeat homology <SA1>
F:221-394/Domain: serum albumin repeat homology <SA2>
F:413-592/Domain: serum albumin repeat homology <SA3>
F:22/Binding site: copper (His) #status predicted
F:99-114,113-124,148-193,192-201,224-270,269-277,289-303,302-313,384-393,416-462,461-472
F:249/Binding site: bilirubin (Lys) #status predicted
F:251/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 32.9%; Score 1245.7; DB 1; Length 609;
Best Local Similarity 39.4%; Pred. No. 9.1e-25;
Matches 234; Conservative 115; Mismatches 232; Indels 13; Gaps 3;

QY 4 KSEVAHR-----PKDLGEENFKALVIAFAQYLQCPFDHVKLVNEVTEFAKT 52
   : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 17 ESRTLHRNRYGSIADSDYCTAETSLADLATIFFAQFVQEAITYEVSVMKDALTAIBK 76
   : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY 53 CVADESAENCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPNERNECFLOHKDNP-N 111
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 77 PTGDEQSACLENQLPAFLLEELCHEILEKYG-LSDDCSQSEGEHNCFLAKKPTPAS 135
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY 112 LPLRVPEVDVMTAFHDNEETFLKYLVEIARRHPYFVAPPELLFPKAKYKAAFTCCQA 171
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 136 IPLFQVPEPVTSCAEYEDRETFMNFYIEIARRHPFYAPTILLWAARYDKIIPSCCKA 195
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY 172 ADKAACLLPKLDELDEGKASSAKQLKCLASLQKGERAFKAWAVARLSORPKAEFAV 231
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 196 ENAVECFQTKAATVTVKELRESSLLNHACAVMKNFGTRTFQAITVTKLSQKFTKNFTBI 255
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY 232 SKLVTLDTKVHTBCCHDGLLECCADDRADLAKYICENQDTSISKLKCECKPLLEKSHCIA 291
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 256 QKLVLDAVHVEHCCRGVDLDCLODEKMTSYICSQDITLSNKTETCCKLTTLERQCII 315
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY 292 EVENDEMPADLPSSAADPVESKDVCKNVAEAKDVFLGMFLMEYARRHPDYSVVLRLAK 351
   ||||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 316 HAENDEKPEGLSPNLRFLGDRDFNQFSSEKKNIFLASFVHEYSRRHPQLAVSVILRVAK 375

```

```

QY 352 TYETTLKCCAAADPHCYAKVDFEKLVPBEPQNLIKONCELPBQLGEYKFNALLVRY 411
   : ||||| : ||| : : : : : ||||| : ||||| : ||||| : ||||| : |||||
Db 376 GYQELLEKCFQTNPLECQDKGEBELQKYIQESQALAKRSCGLFQKLGUYLQNAFLVAY 435
   : ||||| : ||| : : : : : ||||| : ||||| : ||||| : ||||| : |||||
QY 412 TKKYPQVSTPTLTVESVNLGVGSKCKHPKAPKMPCAEDYLSVVLNOLCVLHKTPTVSD 471
   ||| ||||| : : : : : ||||| : ||||| : ||||| : ||||| : |||||
Db 436 TKKAPQLTSSSELMATRKMAATAATCCQLSEDKLLACGEGAADIIGHLCIRHEMTVPNP 495
   ||| ||||| : : : : : ||||| : ||||| : ||||| : ||||| : |||||
QY 472 RVTKCTESLVNRRPCFSALEVDETVYPKFNAAETFTPHADICTLSEKERQIKKQTALVE 531
   : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
Db 496 GVGQCTSSYANRRPCFSLVVDYTVPPAFSDDKFIFHKOLCOAQGVALTQMKQEFLLN 555
   : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
QY 532 LVKHKPKATKQLKAVMDPFAFVEKCKCKADDKTCFAEBGKLVAAQAALGL 585
   ||| ||||| : : : : : ||||| : ||||| : ||||| : ||||| : |||||
Db 556 LVKQKPQITEGLEAVIADFGLLEKCCQGQEQVEVCFABEGQKLSKTRTALGV 609

RESULT 15
ABXL72
74k albumin precursor - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C:Accession: B41682; S02693; A05288
R:Moskaitis, J.E.; Sargent, T.D.; Smith Jr., L.H.; Pastori, R.L.; Schoenberg, D.R.
Mol. Endocrinol. 3, 464-473, 1989
A:Title: Xenopus laevis serum albumin: sequence of the complementary deoxyribonucleic a
e during development.
A:Reference number: A41682; MUID:89313788; PMID:2747653
A:Accession: B41682
A:Molecule type: mRNA
A:Residues: 3-607 <MOS>
A:Cross-references: UNIPROT:P14872; GB:M21442; NID:9213930; PIDN:AAA49637.1; PID:g213931
R:Schorpp, M.; Doeberling, U.; Wagner, U.; Ryffel, G.U.
J. Mol. Biol. 199, 83-93, 1988
A:Title: 5'-flanking and 5'-proximal exon regions of the two Xenopus albumin genes. Del
A:Reference number: S02692; MUID:86172470; PMID:2451026
A:Accession: S02693
A>Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-48 <SCH>
A:Cross-references: EMBL:226826
R:Wolfe, A.P.; Glover, J.F.; Martin, S.C.; Tenniswood, M.P.R.; Williams, J.L.; Tata, J.
Eur. J. Biochem. 146, 489-496, 1985
A:Title: Deinduction of transcription of Xenopus 74-kDa albumin genes and destabilizati
A:Reference number: A05288; MUID:85126974; PMID:3971963
A:Accession: A05288
A:Molecule type: mRNA
A:Residues: 459-502, 'L', 504-557 <WOL>
A:Cross-references: GB:M28276
A>Note: the authors translated the codon TAT for residue 63 as Thr
C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,
mones (weak bonds with these hormones promote their transfer across the membranes), thy
C:Genetics:
A:Introns: 27/1
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: carrier protein; duplication; glycoprotein; metal binding; plasma
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-24/Domain: propeptide #status predicted <PRO>
F:25-607/Product: 74k serum albumin #status predicted <MAT>
F:32-201/Domain: serum albumin repeat homology <SA1>
F:220-393/Domain: serum albumin repeat homology <SA2>
F:412-591/Domain: serum albumin repeat homology <SA3>
F:30/Binding site: copper (His) #status predicted
F:80-88,101-117,116-127,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392
F:256/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 31.9%; Score 1206.6; DB 1; Length 607;
Best Local Similarity 39.3%; Pred. No. 9.7e-24;
Matches 227; Conservative 108; Mismatches 239; Indels 4; Gaps 2;

QY 3 HKSEVAHRPKDLGEENFKALVIAFAQYLQCPFDHVKLVNEVTEFAKTVADESAENC 62
   : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :

```

Db 30 HHHKHIADVYTALTERTFKGLTLAIVSQNLQCSLEELSKLVINEINDFAKSCINDKTPE-C 88

Qy 63 DKSHTLTFGDKLCTVATILRETYGEMADCCAKQOEPRNECFLOHKDDNPNIPLRVPRVDV 122

Db 89 EKPVGTLFFDKLCAADPAVGVNYEWSKECCAKQDPERAQCFKAHRDHBT---SIKPEPEE 145

Qy 123 MCTAFHNEETFLKYLVEYARRHPYFVAPELLFFAKRYKAAFPTECCQAAADKAACLLPKL 182

Db 146 TKLLKHEHPDLLSAFIHEARNHPDLYPPAVLALTQYHKLAHCCBEEDKEKCFSEKM 205

Qy 183 DELRDEGKASSAKQRLKCAQKQGERAFKAWAVARLSORFPKAEFAVSKLVTDLTQVH 242

Db 206 KQLMKQSHSIEDKQHFECWILDNFPEKVLKALNARVSHRYPKAEFKLAHNFTEVTHFI 265

Qy 243 TECCHGDLLECADRADLAKYICENQDSISSKLKECCKEPILLESKHCIAEAVENDEMPADL 302

Db 266 KDCCHDDMFECMTERLEHTCOHKDELSSKLECCNIPLLERTYICVITLNDVDPAEL 325

Qy 303 PSLSAADFVSKDVCNKYAEAKDVFGLGMFLVEYARRHPDYSVLLRLRAKTYETTLEKCCA 362

Db 326 SQPTETEDPHVCEKKAENNEVFLGRYLHAVSRHQBELSEQFLQSAKEYESLLNKCK 385

Qy 363 AADPHECYAKVDFBKFPLVEBPQNLIKQNCLEFBLQGEYKFQNALLVRYTKKVPQVSTPT 422

Db 386 TDNPPECYKDGADRFMNEAKERFAYLKQNCIDLHEHGYLFFENELLIRYTKKMPQVSDET 445

Qy 423 LVEVSRNLKGVSKCKKHPEAKRMPCARBYLSVVLNQLCVLHEKTPVSDRVTKCTESLV 482

Db 446 LIGIAHQWADIGECCCAVPENQPCAEGLDTILIGMKERCQKKTFFNNHVAHCCDTSYS 505

Qy 483 NRRPFSALEVDENYVPEFNAETFTPHADICTLSEKEROIKKOTALVELVGHKPKATKE 542

Db 506 GMRSCFTALGDDEYVPPVTDTHFDKDKICTANDKEKHQKFLVLKLVKSPKLEKN 565

Qy 543 QLKAVMDDEAFVKECKCKADDKETCFABEGKKLVAASQ 580

Db 566 HIDECSAEFLKMVQKCTADEHQCFDTEKPVLTIEHQ 603

RESULT 16

I46986

aluminum - dog (fragment)

C:Species: Canis lupus familiaris (dog)

C:Date: 04-Sep-1997 #sequence\_revision 07-Nov-1997 #text\_change 20-Aug-1999

C:Accession: I46986

R:Spitzauer, S.; Schweiger, C.; Sperr, W.R.; Pandjaitan, B.; Valent, P.; Muhl, S.; Ebner J. Allergy Clin. Immunol. 93, 614-627, 1994

A:Title: Molecular characterization of dog albumin as a cross-reactive allergen.

A:Reference number: I46986; UID:94201492; PMID:7512102

A:Accession: I46986

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-265 <SPI>

A:Cross-references: GB:S72946; NID:g633937; PID:AA830434.1; PID:g633938

C:Superfamily: serum albumin; serum albumin repeat homology

F:7-180/Domain: serum albumin repeat homology <SA2>

Query Match 31.2%; Score 1181.9; DB 2; Length 265;

Best Local Similarity 82.6%; Pred. No. 3.1e-24;

Matches 218; Conservative 23; Mismatches 22; Indels 1; Gaps 1;

Qy 192 SSAQRLKCAQKQGERAFKAWAVARLSORFPKAEFAVSKLVTDLTQVTECHGDL 251

Db 2 SSAKERFKCAQKQGERAFKAWAVARLSORFPKADFAETSKVYTDLTQVHKECCHGDL 61

Qy 252 ECADRADLAKYICENQDSISSKLKECCKEPILLESKHCIAEAVENDEMPADLPSLSAADPVE 311

Db 62 ECADRADLAKYICENQDSISTKLKECCDKVPLEKSQCLAEVERDELPGDLPSSLSAADPVE 121

Qy 312 SKDVCNKYAEAKDVFGLGMFLVEYARRHPDYSVLLRLRAKTYETTLEKCCAAADPHECYA 371

Db 122 DKEVCNKYAEAKDVFGLGMFLVEYARRHPDYSVLLRLRAKTYETTLEKCCATDDPPTCYA 181

Qy 372 KVFDEFKPLVEEPQNLIKQNCLEFBLQGEYKFQNALLVRYTKKVPQVSTPTL-VEVSRNL 430

Db 182 KVLDEFKPLVDEPQNLVKTNCLEFKEGLGFGQNALLVRYTKKAPQVSTPTLVVVEVRKL 241

Qy 431 GKVSGKCKKHPEAKRMPCARBYLS 454

Db 242 KGVTGTCCKKPESERMSCADDFLS 265

RESULT 17

ABXL68

68K serum albumin precursor - African clawed frog

C:Species: Xenopus laevis (African clawed frog)

C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004

C:Accession: A41682; S02692

R:Moakatis, J.E.; Sargent, T.D.; Smith Jr., L.H.; Pastori, R.L.; Schoenberg, D.R. Mol. Endocrinol. 3, 464-473, 1989

A:Title: Xenopus laevis serum albumin: sequence of the complementary deoxyribonucleic acid during development.

A:Reference number: A41682; UID:89313788; PMID:2747653

A:Accession: A41682

A:Molecule type: mRNA

A:Residues: 1-608 <MOS>

A:Cross-references: UNIPROT:P08759; GB:M18350

R:Schorpp, M.; Doebebeling, U.; Wagner, U.; Ryffel, G.U. J. Mol. Biol. 199, 83-93, 1988

A:Title: 5'-flanking and 5'-proximal exon regions of the two Xenopus albumin genes. Deleted exon 5' is transcribed but not translated.

A:Reference number: S02692; UID:88172470; PMID:2451026

A:Accession: S02692

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-48 <SCH>

A:Cross-references: EMBL:Z26825

C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper, moneis (weak bonds with these hormones promote their transfer across the membranes), thyronines, and thyroglobulin.

C:Genetics: 27/1

A:Introns: 27/1

C:Superfamily: serum albumin; serum albumin repeat homology

C:Keywords: carrier protein; duplication; metal binding; plasma

F:1-18/Domain: signal sequence #status predicted <SIG>

F:19-24/Domain: propeptide #status predicted <PRO>

F:25-608/Product: 68K serum albumin #status predicted <MAP>

F:32-202/Domain: serum albumin repeat homology <SA1>

F:221-394/Domain: serum albumin repeat homology <SA2>

F:413-592/Domain: serum albumin repeat homology <SA3>

F:30/Binding site: copper (His) #status predicted

F:80-89,102-118,117-128,148-193,192-201,224-270,269-277,289-303,302-313,340-385,384-393,389-390,402-403,404-405,406-407,408-409,410-411,412-413,414-415,416-417,418-419,420-421,422-423,424-425,426-427,428-429,430-431,432-433,434-435,436-437,438-439,440-441,442-443,444-445,446-447,448-449,450-451,452-453,454-455,456-457,458-459,460-461,462-463,464-465,466-467,468-469,470-471,472-473,474-475,476-477,478-479,480-481,482-483,484-485,486-487,488-489,490-491,492-493,494-495,496-497,498-499,500-501,502-503,504-505,506-507,508-509,510-511,512-513,514-515,516-517,518-519,520-521,522-523,524-525,526-527,528-529,530-531,532-533,534-535,536-537,538-539,540-541,542-543,544-545,546-547,548-549,550-551,552-553,554-555,556-557,558-559,560-561,562-563,564-565,566-567,568-569,570-571,572-573,574-575,576-577,578-579,580-581,582-583,584-585,586-587,588-589,590-591,592-593,594-595,596-597,598-599,600-601,602-603,604-605,606-607,608-609,610-611,612-613,614-615,616-617,618-619,620-621,622-623,624-625,626-627,628-629,630-631,632-633,634-635,636-637,638-639,640-641,642-643,644-645,646-647,648-649,650-651,652-653,654-655,656-657,658-659,660-661,662-663,664-665,666-667,668-669,670-671,672-673,674-675,676-677,678-679,680-681,682-683,684-685,686-687,688-689,690-691,692-693,694-695,696-697,698-699,700-701,702-703,704-705,706-707,708-709,710-711,712-713,714-715,716-717,718-719,720-721,722-723,724-725,726-727,728-729,730-731,732-733,734-735,736-737,738-739,740-741,742-743,744-745,746-747,748-749,750-751,752-753,754-755,756-757,758-759,760-761,762-763,764-765,766-767,768-769,770-771,772-773,774-775,776-777,778-779,780-781,782-783,784-785,786-787,788-789,790-791,792-793,794-795,796-797,798-799,800-801,802-803,804-805,806-807,808-809,810-811,812-813,814-815,816-817,818-819,820-821,822-823,824-825,826-827,828-829,830-831,832-833,834-835,836-837,838-839,840-841,842-843,844-845,846-847,848-849,850-851,852-853,854-855,856-857,858-859,860-861,862-863,864-865,866-867,868-869,870-871,872-873,874-875,876-877,878-879,880-881,882-883,884-885,886-887,888-889,890-891,892-893,894-895,896-897,898-899,900-901,902-903,904-905,906-907,908-909,910-911,912-913,914-915,916-917,918-919,920-921,922-923,924-925,926-927,928-929,930-931,932-933,934-935,936-937,938-939,940-941,942-943,944-945,946-947,948-949,950-951,952-953,954-955,956-957,958-959,960-961,962-963,964-965,966-967,968-969,970-971,972-973,974-975,976-977,978-979,980-981,982-983,984-985,986-987,988-989,990-991,992-993,994-995,996-997,998-999,1000-1001,1002-1003,1004-1005,1006-1007,1008-1009,1010-1011,1012-1013,1014-1015,1016-1017,1018-1019,1020-1021,1022-1023,1024-1025,1026-1027,1028-1029,1030-1031,1032-1033,1034-1035,1036-1037,1038-1039,1040-1041,1042-1043,1044-1045,1046-1047,1048-1049,1050-1051,1052-1053,1054-1055,1056-1057,1058-1059,1060-1061,1062-1063,1064-1065,1066-1067,1068-1069,1070-1071,1072-1073,1074-1075,1076-1077,1078-1079,1080-1081,1082-1083,1084-1085,1086-1087,1088-1089,1090-1091,1092-1093,1094-1095,1096-1097,1098-1099,1100-1101,1102-1103,1104-1105,1106-1107,1108-1109,1110-1111,1112-1113,1114-1115,1116-1117,1118-1119,1120-1121,1122-1123,1124-1125,1126-1127,1128-1129,1130-1131,1132-1133,1134-1135,1136-1137,1138-1139,1140-1141,1142-1143,1144-1145,1146-1147,1148-1149,1150-1151,1152-1153,1154-1155,1156-1157,1158-1159,1160-1161,1162-1163,1164-1165,1166-1167,1168-1169,1170-1171,1172-1173,1174-1175,1176-1177,1178-1179,1180-1181,1182-1183,1184-1185,1186-1187,1188-1189,1190-1191,1192-1193,1194-1195,1196-1197,1198-1199,1200-1201,1202-1203,1204-1205,1206-1207,1208-1209,1210-1211,1212-1213,1214-1215,1216-1217,1218-1219,1220-1221,1222-1223,1224-1225,1226-1227,1228-1229,1230-1231,1232-1233,1234-1235,1236-1237,1238-1239,1240-1241,1242-1243,1244-1245,1246-1247,1248-1249,1250-1251,1252-1253,1254-1255,1256-1257,1258-1259,1260-1261,1262-1263,1264-1265,1266-1267,1268-1269,1270-1271,1272-1273,1274-1275,1276-1277,1278-1279,1280-1281,1282-1283,1284-1285,1286-1287,1288-1289,1290-1291,1292-1293,1294-1295,1296-1297,1298-1299,1300-1301,1302-1303,1304-1305,1306-1307,1308-1309,1310-1311,1312-1313,1314-1315,1316-1317,1318-1319,1320-1321,1322-1323,1324-1325,1326-1327,1328-1329,1330-1331,1332-1333,1334-1335,1336-1337,1338-1339,1340-1341,1342-1343,1344-1345,1346-1347,1348-1349,1350-1351,1352-1353,1354-1355,1356-1357,1358-1359,1360-1361,1362-1363,1364-1365,1366-1367,1368-1369,1370-1371,1372-1373,1374-1375,1376-1377,1378-1379,1380-1381,1382-1383,1384-1385,1386-1387,1388-1389,1390-1391,1392-1393,1394-1395,1396-1397,1398-1399,1400-1401,1402-1403,1404-1405,1406-1407,1408-1409,1410-1411,1412-1413,1414-1415,1416-1417,1418-1419,1420-1421,1422-1423,1424-1425,1426-1427,1428-1429,1430-1431,1432-1433,1434-1435,1436-1437,1438-1439,1440-1441,1442-1443,1444-1445,1446-1447,1448-1449,1450-1451,1452-1453,1454-1455,1456-1457,1458-1459,1460-1461,1462-1463,1464-1465,1466-1467,1468-1469,1470-1471,1472-1473,1474-1475,1476-1477,1478-1479,1480-1481,1482-1483,1484-1485,1486-1487,1488-1489,1490-1491,1492-1493,1494-1495,1496-1497,1498-1499,1500-1501,1502-1503,1504-1505,1506-1507,1508-1509,1510-1511,1512-1513,1514-1515,1516-1517,1518-1519,1520-1521,1522-1523,1524-1525,1526-1527,1528-1529,1530-1531,1532-1533,1534-1535,1536-1537,1538-1539,1540-1541,1542-1543,1544-1545,1546-1547,1548-1549,1550-1551,1552-1553,1554-1555,1556-1557,1558-1559,1560-1561,1562-1563,1564-1565,1566-1567,1568-1569,1570-1571,1572-1573,1574-1575,1576-1577,1578-1579,1580-1581,1582-1583,1584-1585,1586-1587,1588-1589,1590-1591,1592-1593,1594-1595,1596-1597,1598-1599,1600-1601,1602-1603,1604-1605,1606-1607,1608-1609,1610-1611,1612-1613,1614-1615,1616-1617,1618-1619,1620-1621,1622-1623,1624-1625,1626-1627,1628-1629,1630-1631,1632-1633,1634-1635,1636-1637,1638-1639,1640-1641,1642-1643,1644-1645,1646-1647,1648-1649,1650-1651,1652-1653,1654-1655,1656-1657,1658-1659,1660-1661,1662-1663,1664-1665,1666-1667,1668-1669,1670-1671,1672-1673,1674-1675,1676-1677,1678-1679,1680-1681,1682-1683,1684-1685,1686-1687,1688-1689,1690-1691,1692-1693,1694-1695,1696-1697,1698-1699,1700-1701,1702-1703,1704-1705,1706-1707,1708-1709,1710-1711,1712-1713,1714-1715,1716-1717,1718-1719,1720-1721,1722-1723,1724-1725,1726-1727,1728-1729,1730-1731,1732-1733,1734-1735,1736-1737,1738-1739,1740-1741,1742-1743,1744-1745,1746-1747,1748-1749,1750-1751,1752-1753,1754-1755,1756-1757,1758-1759,1760-1761,1762-1763,1764-1765,1766-1767,1768-1769,1770-1771,1772-1773,1774-1775,1776-1777,1778-1779,1780-1781,1782-1783,1784-1785,1786-1787,1788-1789,1790-1791,1792-1793,1794-1795,1796-1797,1798-1799,1800-1801,1802-1803,1804-1805,1806-1807,1808-1809,1810-1811,1812-1813,1814-1815,1816-1817,1818-1819,1820-1821,1822-1823,1824-1825,1826-1827,1828-1829,1830-1831,1832-1833,1834-1835,1836-1837,1838-1839,1840-1841,1842-1843,1844-1845,1846-1847,1848-1849,1850-1851,1852-1853,1854-1855,1856-1857,1858-1859,1860-1861,1862-1863,1864-1865,1866-1867,1868-1869,1870-1871,1872-1873,1874-1875,1876-1877,1878-1879,1880-1881,1882-1883,1884-1885,1886-1887,1888-1889,1890-1891,1892-1893,1894-1895,1896-1897,1898-1899,1900-1901,1902-1903,1904-1905,1906-1907,1908-1909,1910-1911,1912-1913,1914-1915,1916-1917,1918-1919,1920-1921,1922-1923,1924-1925,1926-1927,1928-1929,1930-1931,1932-1933,1934-1935,1936-1937,1938-1939,1940-1941,1942-1943,1944-1945,1946-1947,1948-1949,1950-1951,1952-1953,1954-1955,1956-1957,1958-1959,1960-1961,1962-1963,1964-1965,1966-1967,1968-1969,1970-1971,1972-1973,1974-1975,1976-1977,1978-1979,1980-1981,1982-1983,1984-1985,1986-1987,1988-1989,1990-1991,1992-1993,1994-1995,1996-1997,1998-1999,2000-2001,2002-2003,2004-2005,2006-2007,2008-2009,2010-2011,2012-2013,2014-2015,2016-2017,2018-2019,2020-2021,2022-2023,2024-2025,2026-2027,2028-2029,2030-2031,2032-2033,2034-2035,2036-2037,2038-2039,2040-2041,2042-2043,2044-2045,2046-2047,2048-2049,2050-2051,2052-2053,2054-2055,2056-2057,2058-2059,2060-2061,2062-2063,2064-2065,2066-2067,2068-2069,2070-2071,2072-2073,2074-2075,2076-2077,2078-2079,2080-2081,2082-2083,2084-2085,2086-2087,2088-2089,2090-2091,2092-2093,2094-2095,2096-2097,2098-2099,2100-2101,2102-2103,2104-2105,2106-2107,2108-2109,2110-2111,2112-2113,2114-2115,2116-2117,2118-2119,2120-2121,2122-2123,2124-2125,2126-2127,2128-2129,2130-2131,2132-2133,2134-2135,2136-2137,2138-2139,2140-2141,2142-2143,2144-2145,2146-2147,2148-2149,2150-2151,2152-2153,2154-2155,2156-2157,2158-2159,2160-2161,2162-2163,2164-2165,2166-2167,2168-2169,2170-2171,2172-2173,2174-2175,2176-2177,2178-2179,2180-2181,2182-2183,2184-2185,2186-2187,2188-2189,2190-2191,2192-2193,2194-2195,2196-2197,2198-2199,2200-2201,2202-2203,2204-2205,2206-2207,2208-2209,2210-2211,2212-2213,2214-2215,2216-2217,2218-2219,2220-2221,2222-2223,2224-2225,2226-2227,2228-2229,2230-2231,2232-2233,2234-2235,2236-2237,2238-2239,2240-2241,2242-2243,2244-2245,2246-2247,2248-2249,2250-2251,2252-2253,2254-2255,2256-2257,2258-2259,2260-2261,2262-2263,2264-2265,2266-2267,2268-2269,2270-2271,2272-2273,2274-2275,2276-2277,2278-2279,2280-2281,2282-2283,2284-2285,2286-2287,2288-2289,2290-2291,2292-2293,2294-2295,2296-2297,2298-2299,2300-2301,2302-2303,2304-2305,2306-2307,2308-2309,2310-2311,2312-2313,2314-2315,2316-2317,2318-2319,2320-2321,2322-2323,2324-2325,2326-2327,2328-2329,2330-2331,2332-2333,2334-2335,2336-2337,2338-2339,2340-2341,2342-2343,2344-2345,2346-2347,2348-2349,2350-2351,2352-2353,2354-2355,2356-2357,2358-2359,2360-2361,2362-2363,2364-2365,2366-2367,2368-2369,2370-2371,2372-2373,2374-2375,2376-2377,2378-2379,2380-2381,2382-2383,2384-2385,2386-2387,2388-2389,2390-2391,2392-2393,2394-2395,2396-2397,2398-2399,2400-2401,2402-2403,2404-2405,2406-2407,2408-2409,2410-2411,2412-2413,2414-2415,2416-2417,2418-2419,2420-2421,2422-24

```
Db 325 ELSPITEFTEDPHVCEKYAENKESFLERIISPMOSQETPELSEQLQSAKEYSLLNKC 384
Qy 361 CAADPHCYAKVDFEKPPLVEEPQNLIKQNCLEFPEQLGEYKFNQALVRYTKVPQVST 420
Db 385 CFSQNPPECYKDGADRFNNEAKERAYLKQNCIDLHEGVEYLFENELLIRYTKWMPQVSD 444
Qy 421 PTLVEVSRLGKVGSKCCKHPEAKMPCAEDYLSVLNQLCVLHFKTPVSDRVTKCTPES 480
Db 445 ETLIGIAHQMADIGEHCCAVPENQMPCAEGLDTLIGKMCERQKKTINNVAHCCYDS 504
Qy 481 LVNRPFCFALEVDYTPKPEFNAETFTFHADICTLSEKERQIKKOTALVELVGHKPKAT 540
Db 505 YSGMRSCFTALGPDSDDYPPVPTDFTFHDKICTANDKEQHOKKQFLVXLKIRVSPKLE 564
Qy 541 KEQLKAVMDDDFAAFYEKCKADDKETCFAEKGKLVAAASQ 580
Db 565 KNHIDWLEFLKMWQKCTADEHQPCDTEKPVLIENHCQ 604

RESULT 18
PPMS
alpha-fetoprotein precursor - mouse
N:Alternate names: AFP; alpha-1-fetoprotein precursor; alpha-fetoglobulin precursor
C:Species: Mus musculus (house mouse)
C>Date: 01-Sep-1981 #sequence revision 28-May-1986 #text_change 09-Jul-2004
C:Accession: A93254; A93305; A03235
R:Law, S.W.; Dugaiczky, A.
Nature 291, 201-205, 1981
A>Title: Homology between the primary structure of alpha-fetoprotein, deduced from a com
A:Reference number: A93254; MUID:81197641; PMID:6164927
A:Accession: A93254
A:Molecule type: mRNA
A:Residues: 1-597; 'E', 599-605 <LAW>
A:Cross-references: UNIPROT:P02772; GB:V00743
R:Mirghetti, P.P.; Law, S.W.; Dugaiczky, A.
Mol. Biol. Evol. 2, 347-358, 1985
A>Title: The rate of molecular evolution of alpha-fetoprotein approaches that of pseudog
A:Reference number: A93055; MUID:88216123; PMID:2452956
A:Contents: annotation; revision to residue 598
A>Note: residue 598, reported in reference A93254 as Glu, should have been reported as L
R:Gorin, M.B.; Cooper, D.L.; Eiferman, F.; van de Rijn, P.; Tilghman, S.M.
J. Biol. Chem. 256, 1954-1959, 1981
A>Title: The evolution of alpha-fetoprotein and albumin: I. A comparison of the primary
A:Reference number: A93305; MUID:81117287; PMID:6161929
A:Accession: A92305
A:Molecule type: mRNA
A:Residues: 15-532; 'RAKL', 538-605 <GOR>
A:Cross-references: GB:M16381
A>Note: the beginning of the mature protein was placed at residue 21
R:Eiferman, F.A.; Young, P.R.; Scott, R.W.; Tilghman, S.M.
Nature 294, 713-718, 1981
A>Title: Intragenic amplification and divergence in the mouse alpha-fetoprotein gene.
A:Reference number: A93271; MUID:82103990; PMID:6172714
A:Contents: annotation; exon-intron junctions
C:Comment: Mouse AFP has two carbohydrate chains and was found to bind estrogens and cop
C:Genetics:
A:Map position: 5
A:Introns: 29/1; 46/2; 86/3; 157/2; 201/3; 234/2; 277/3; 349/2; 393/3; 426/2; 472/3; 547
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: embryo; fetus; globulin; glycoprotein; metal binding; plasma
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-605/Product: alpha-fetoprotein #status predicted <MPT>
F:29-198/Domain: serum albumin repeat homology <SA1>
F:217-390/Domain: serum albumin repeat homology <SA2>
F:409-588/Domain: serum albumin repeat homology <SA3>
F:22/Binding site: copper (His) #status predicted
F:95-110,109-120,144-189,188-197,220-266,265-273,285-299,298-309,380-389,412-458,457-468
F:247,498/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 28.7%; Score 1086.8; DB 1; Length 605;
Best Local Similarity 35.6%; Pred. No. 1.4e-20;
Matches 206; Conservative 121; Mismatches 230; Indels 22; Gaps 5;
```

```
Qy 16 BENFKALVLIAPQYLQCCPPEDHVKLVNEVTEFAKTCVADESANCKSLHTLPGDKLC 75
Db 40 EKNVLSIATITFTQVPEATEEE---VNKMTSDVLAAMKKNKSGDGLSLSQSFVLEIC 95
Qy 76 TVATLRETYGMAKCCAKOBERNECFLOHKDDNP-NLPLRVLRPEVDVMCTAFTHDNEETF 134
Db 96 HETELSNKYG-LSGCCSQSGVERHQCLLARKKTAPASVPPFPFPEPAESCKAHENRAVF 154
Qy 135 LKLYLIEIARHPHYFAPELLFAKRYKAAFTCCQAADKAAACLLPKLDELDEGKASSA 194
Db 155 MNRFIYVSRNRPMPYAPAILSLAAQYDKVVLVACKADNKECFQTK-----RASIA 206
Qy 195 KORLK-----CASLQKFGERAFAKAWARLQRPFPKAFSAFVSKLVTLTKVHTGCC 246
Db 207 KELREGSLNHEHVCVIRKFGSRNLQATTTIKLSQKLTEANFTEIQLKALDVVAHTHECC 266
Qy 247 HGDILLECADRADLAKYICENQDSITSSKLEKCEKFLLEKSHCIAEVENDEMPADLPSLA 306
Db 267 QGNSLELQDGEKVTYICSQNILSSKIAECCKLPMIQLGFCIIHAENGKPEGLSLNP 326
Qy 307 ADFVESKDVCKNYAEAKDVFLGMFLYEHYARHPDYSVVLLRLAKTYTTTLEKCCAAADP 366
Db 327 SQFLGDRNFAQFSSEKIMFMASFLHEYSRTHPNLPVSVILRIAKTYQEIILEKCSQSNL 386
Qy 367 HECYAKVDFEKPPLVEEPQNLIKQNCLEFPEQLGEYKFNQALVRYTKVPQVSTFLVEV 426
Db 387 PGCDNLFEELQKHIEESQALSQSCALYQTLGDKLQNLFLIGYTRKAPQLTSAELIDL 446
Qy 427 SRNLGKVGSKCKHPEAKRMPCAEDYLSVLNQLCVLHFKTPVSDRVTKCTPESLVNRRP 486
Db 447 TGMVSTASTCCQLSEKWSGCGGEMADFTGHLICIRNEASFPVNSGISHCCNSSVSNRL 506
Qy 487 CFSALEVDVETVPKFEFNAETFTFHADICTLSEKERQIKKOTALVELVHKPKATKEQLKA 546
Db 507 CITSFLRDETYAPPFSEDKFIHFKDLQCAQKALQTMQELLINLVQKPELTTEEQLAA 566
Qy 547 VMDPFAFVEKCKKADDKETCFAEKGKLVAAASQALGL 585
Db 567 VTADPSGLLEKCKCAQDOEVCFTBEGPKLISKTRDALGV 605
```

## RESULT 19

## FPRT

```
alpha-fetoprotein precursor - rat
N:Alternate names: AFP precursor; alpha-1-fetoprotein precursor; alpha-fetoglobulin pre
C:Species: Rattus norvegicus (Norway rat)
C>Date: 08-Oct-1981 #sequence revision 28-May-1986 #text change 09-Jul-2004
C:Accession: A93561; A93876; JN0247; I53048; I55253; I55217; I55218; I69969; I58260; A03
R:Turcotte, B.; Guertin, M.; Chevreton, M.; Belanger, L.
Nucleic Acids Res. 13, 2387-2398, 1985
A>Title: Rat alpha 1-fetoprotein messenger RNA: 5'-end sequence and glucocorticoid-suppl
A:Reference number: A93561; MUID:85215621; PMID:2582363
A:Accession: A93561
A:Molecule type: mRNA
A:Residues: 1-106 <TUR>
A:Cross-references: UNIPROT:P02773; GB:X02361
R:Jagodzinski, L.L.; Sargent, T.D.; Yang, M.; Glackin, C.; Bonner, J.
Proc. Natl. Acad. Sci. U.S.A. 78, 3521-3525, 1981
A>Title: Sequence homology between RNAs encoding rat alpha-fetoprotein and rat serum al
A:Reference number: A93876; MUID:81273091; PMID:6167988
A:Accession: A93876
A:Molecule type: mRNA
A:Residues: 91-611 <JAG>
A:Cross-references: GB:V01254; NID:956781; PIDN:CAA24567.1; PID:g809077
R:Watanabe, T.; Jimenez-Molina, J.L.; Chou, J.Y.
Biochem. Biophys. Res. Commun. 185, 648-656, 1992
A>Title: Characterization of a rat variant alpha-fetoprotein.
A:Reference number: JN0247; MUID:92304289; PMID:1376990
A:Accession: JN0247
A:Molecule type: mRNA
A:Residues: 287-611 <WAT>
A:Experimental source: liver
```















A;Introns: 20/1; 43/2; 87/3; 158/2; 202/3; 234/2; 277/3; 345/2; 388/3; 421/2; 465/3  
C;Superfamily: serum albumin; serum albumin repeat homology  
C;Keywords: actin binding; duplication; globulin; glycoprotein; liver; plasma; polymorph  
F;1-16/Domain: signal sequence #status predicted <SIG>  
F;17-476/Product: vitamin D-binding protein #status experimental <MPT>  
F;26-199/Domain: serum albumin repeat homology <SA1>  
F;217-385/Domain: serum albumin repeat homology <SA2>  
F;404-476/Domain: serum albumin repeat homology #status atypical <SA3>  
F;29-75,74-83,96-112,111-122,145-190,189-198,220-266,265-273,286-300,299-311,335-376,379  
F;288/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 11.1%; Score 421; DB 1; Length 476;  
Best Local Similarity 22.6%; Pred. No. 0.0025;  
Matches 123; Conservative 82; Mismatches 190; Indels 150; Gaps 20;

QY 1 DAHSEVAHRFKDGLGEENFKALVIAFAQYLQCCPFEDHVLVNEVTEFAKTCVADESAE 60  
DB 22 DYEDKKVCQELSTGLKDDFRSLILYSRKPSSFTFEQVSQVLKEVVSLECCBAGADP 81  
QY 61 NCDKSLHTLFGDKLCTVATLRETGYEMADCAKQBPENECF--LOHKDDNPNIPLRVLP 118  
DB 82 NCYDTRTSELSIKSCSDAPPPVHPGTSECTCKEGLERKLCMAALSHQPQ--EFPAYVEP 139  
QY 119 EVDVCTAFHNEETFLKYLVEIARRHPYFAPPELLFPAKRYKAAFTCCQADKAACL 178  
DB 140 TNDEICEAFRDKPGFADQFLFEYSNRYGOAPLPLLVGYTKSYLSMWGSCCTSAKPTVCF 199  
QY 179 LPKLDLDEGRKASSAKORLK-----CASLQKFGFRAFKAVARLSQRFP 224  
DB 200 L-----KERLQKQLSLTTMSNRVCSQVAAVGKESRMSHLIKLAQKVP 244  
QY 225 KAFAEVSGLVTDLPKVTCHQCHGLLEC--ADDRADLAKYICENQDSTSSKLKECC---- 279  
DB 245 TANLEDVLPALDELFEILSRCKSTSEDCMARELPHTLTKICGNLSKKNKSPFEECCYETI 304  
QY 280 -----EKLPLEKSHCIAEVENDEMPA--DLPSLAADFVESKDVC--KNVAEKD 324  
DB 305 PMGIFMCSYFMPATBFL-----QUPAIKLPT-----SKDLGQSQTQAMD 344  
QY 375 VFLGFLFVEYARRH--PDYSVLLRLAKTYETTLKCCAAADPHECVAKVDFEKLVE 382  
DB 345 ----QYTFELSRTOVPE---VFLSKVLDTLTLRECDTQDSVSCFS-----TQ 388  
QY 383 EPQNLIKONCELFEOLG-----EYKFONALLVRYTKVPQVSTPTLVEVRN 429  
DB 389 SP--LMKQLTSFIEKGQEMCADYSENTFTYK--KLAERLRTKMPNASPEELADM-- 441  
QY 430 LKGVSKCKKHPKAPCAEDYLSVLNLQCLVHLKPTVSDRVTKCTTESLVNRRPCFS 489  
DB 442 -----VAKH-----SDFASKCCS---INSPPRYC 462  
QY 490 ALEVD 494  
DB 463 SSQID 467

RESULT 28  
VTHUD  
vitamin D-binding protein precursor [validated] - human  
N;Alternate names: DBP; Gc-globulin; group-specific component  
C;Species: Homo sapiens (man)  
C;Date: 28-May-1986 #sequence revision 28-May-1986 #text change 09-Jul-2004  
C;Accession: A94076; A46759; A29096; A92765; S39787; A24066; A90427; A03237  
R;Yang, F.; Brune, J.L.; Naylor, S.L.; Cupples, R.L.; Naberhaus, K.H.; Bowman, B.H.  
Proc. Natl. Acad. Sci. U.S.A. 82, 7994-7998, 1985  
A;Title: Human group-specific component (Gc) is a member of the albumin family.  
A;Reference number: A94076; MUID:86068030; PMID:2415977  
A;Accession: A94076  
A;Molecule type: mRNA  
A;Residues: 1-474 <YANI>  
A;Cross-references: UNIPROT:P02774; GB:X03178; GB:M11321; NID:g31675; PIDN:CAA26938.1; F  
A;Experimental source: allele Gc2  
R;Witke, W.F.; Gibbs, P.E.M.; Zielinski, R.; Yang, F.; Bowman, B.H.; Dugaiczkyk, A.

Genomics 16, 751-754, 1993  
A;Title: Complete structure of the human Gc gene: differences and similarities between m  
A;Reference number: A46759; MUID:93315171; PMID:8325650  
A;Accession: A46759  
A;Status: nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-431, 'E', 433-435, 'T', 437-474 <WIT>  
A;Cross-references: GB:L10641; NID:g340281; PIDN:AAA61704.1; PID:g639896  
R;Yang, F.; Naberhaus, K.H.; Adrian, G.S.; Gardella, J.M.; Brissenden, J.E.; Bowman, B.H.  
Gene 54, 285-290, 1987  
A;Title: The vitamin D-binding protein gene contains conserved nucleotide sequences that  
A;Reference number: A29096; MUID:88005794; PMID:2958390  
A;Accession: A29096  
A;Status: translation not shown  
A;Molecule type: DNA  
A;Residues: 1-19 <YAN2>  
A;Cross-references: GB:M17156; NID:g181489; PIDN:AAA19662.1; PID:g463096  
R;Cooke, N.E.; David, E.V.  
J. Clin. Invest. 76, 2420-2424, 1985  
A;Title: Serum vitamin D-binding protein is a third member of the albumin and alpha fet  
A;Reference number: A92765; MUID:86086396; PMID:2416779  
A;Accession: A92765  
A;Molecule type: mRNA  
A;Residues: 1-167, 'E', 169-326, 'R', 328-431, 'E', 433-435, 'T', 437-474 <COO>  
A;Cross-references: GB:M12634; NID:g181481; PIDN:AAA52173.1; PID:g181482  
A;Experimental source: allele Gc1  
R;Braun, A.; Kofler, A.; Morawietz, S.; Cleve, H.  
Biochim. Biophys. Acta 871, 189-198, 1986  
A;Title: Sequence and organization of the human vitamin D-binding protein gene.  
A;Reference number: S39787; MUID:94092730; PMID:7505619  
A;Accession: S39787  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-431, 'E', 433-435, 'T', 437-474 <BRA>  
A;Cross-references: GB:S67480; NID:g455967; PIDN:AA29423.1; PID:g455970  
R;Schoengen, F.; Meta-Boutigue, M.H.; Jolles, J.; Constans, J.; Jolles, P.  
Biochim. Biophys. Acta 871, 189-198, 1986  
A;Title: Complete amino acid sequence of human vitamin D-binding protein (group-specifi  
A;Reference number: A24066; MUID:86216223; PMID:2423133  
A;Accession: A24066  
A;Molecule type: protein  
A;Residues: 17-474 <SCH>  
R;Svasti, J.; Kurosky, A.; Bennett, A.; Bowman, B.H.  
Biochemistry 18, 1611-1617, 1979  
A;Title: Molecular basis for the three major forms of human serum vitamin D binding pro  
A;Reference number: A90427; MUID:79145448; PMID:218624  
A;Accession: A90427  
A;Molecule type: protein  
A;Residues: 17, 'Q', 19-21, 'N', 23-36, 'XXX', 40-41, 472-474 <SVA>  
C;Comment: DBP is a multifunctional protein found in plasma, ascitic fluid, cerebrosp  
nts polymerization of actin by binding its monomers. DBP associates with membrane-bound  
C;Comment: Over 80 variants of human DBP have been identified. The three most common al  
C;Genetics:  
A;Gene: GDB:GC  
A;Cross-references: GDB:119263; OMIM:139200  
A;Map position: 4q12-q13  
A;Introns: 20/1; 43/2; 87/3; 158/2; 202/3; 234/2; 277/3; 345/2; 388/3; 421/2; 465/3  
C;Superfamily: serum albumin; serum albumin repeat homology  
C;Keywords: actin binding; duplication; globulin; glycoprotein; liver; plasma; polymorpi  
F;1-16/Domain: signal sequence #status predicted <SIG>  
F;17-474/Product: vitamin D-binding protein (allele Gc2) #status experimental <NAL2>  
F;17-431, 'E', 433-435, 'T', 437-474/Product: vitamin D-binding protein (allele Gc1) #statu  
F;26-199/Domain: serum albumin repeat homology <SA1>  
F;217-385/Domain: serum albumin repeat homology <SA2>  
F;404-474/Domain: serum albumin repeat homology #status atypical <SA3>  
F;29-75,74-83,96-112,111-122,145-190,189-198,220-266,265-273,286-300,299-311,335-376,37  
F;288/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 11.0%; Score 417.5; DB 1; Length 474;  
Best Local Similarity 19.5%; Pred. No. 0.003;  
Matches 123; Conservative 91; Mismatches 183; Indels 235; Gaps 22;



QY 76 -----TVATLRETYGEMADCCAKQEPERNECFLOHKDDN----- 109  
Db 272 DITLHNHNSALAAVVPTLTFLAGDCC-----NLSDNEGQVMSAGWTEE 318  
QY 110 -----PNLPLVRPEVDVMCTAFHDNEETFLKK 137  
Db 319 PVALIORMLFRTVLHMSVDISMAEVPENLRKNTLTLRAALKI-----RTFLEK 369  
QY 138 YLYEIARRHPYFAPPELLFFAKRYKA-----APTECCQAADKAACLPLKDELDRDE 188  
Db 370 -----QDPD--FAPRQKTLQEVQDDFVFSKYCH-----RVLLLPELLE--- 406  
QY 189 GKASAKORLKCA-----SLQKGERAFKA-----WAVARLSQR 222  
Db 407 -----GVLOQLILCLOSAASNPFSSQAMDLVQEFIQHGHFLFETAVLQMEWLVR--DG 460  
QY 223 PP-----KAEFAVSKLVDTLTKVHTE----- 244  
Db 461 VPPEASGHLKALINNMKIMSTVKVKYSEQLHSHMCTKRHRRCBYSHFMHHRDLGILL 520  
QY 245 ----- 244  
Db 521 VSAFNQVSKNPFETADGDVYPERCCCCIAVCAHQCLRLQLQAASLSSTCVQLSGVQNI 580  
QY 245 ---CHGD-----LLE-----C 253  
Db 581 GICCCMDPKSVIVPLLHAFKLPALKSCQHILNKLILDLQLGAEIPQKTKAACNIC 640  
QY 254 ADDRADIAX-----YICENQDSIS----- 273  
Db 641 TVDSQOLAKLEETLOGSSYNAPFSSGLSSPSYRFOGILPSSGSDLLMKWDALAEYQNFV 700  
QY 274 -----KLKCECKPILLEKS-----HC----- 289  
Db 701 FEEDRLQSVQIANHICSLIQGNVVQWKLNCIFNPVLQGVVELAHHCQQLSISSAQTH 760  
QY 290 IAEVENDPAD-----LPSLAADFV----- 310  
Db 761 VCSHENQCLPQVQLIYKLTPLTLKSRVIRDLFLSCNGVNIILNLYLDGIRNHSKAF 820  
QY 311 -----ESKOVCKNYAEAKVP 326  
Db 821 ETLIISLGEQKDSIPGIDGLDEQKELSSNVGLYNQAYSDQSLSKFYARKDAY 880  
QY 327 -----LGMFLYEYARRHPDYSVVLLRLAKTYETLEKC-----C 361  
Db 881 PKKRKSVVQDIHISTINFL-----CVAFLCVSKEAESDRESANDESDTSGYDS 929  
QY 362 AAADP-----HE-----CY 370  
Db 930 TASEPLOHMLPCFSLLESVLPSPRMHQAADVMSMCRWIYMLSPFRKQFYRLGGFQVCH 989  
QY 371 AKVF-----DEFKPLVEE-----PQNLIKQNCELFEQOLGEYKQFN 405  
Db 990 KLIFMVIQDLFRNPEEQKRGKEDRTMNENQDLNRSIQPEITVKED----- 1035  
QY 406 ALLVRYTKKVPQVSTPTLVEVSR-----NLGKVGSKCKCKHPEAKRMPCAE----- 450  
Db 1036 ---LLSLTVKI-----DPTPTLSLKKSAADSLGKLESEHLSINVEQIPAVEAVPEETKVF 1089  
QY 451 -----DYLGSVLN 458  
Db 1090 MSRESETLQIGIRLLEALLAICLHGTRASQOKLELELPQNLSVETILLEMEDHLS--- 1145  
QY 459 QLCVLHEKTPVSD-----RV----- 473  
Db 1146 -----KSKVTETELAKPLFDALLRVALGNHNSADFEHDDAMTEKSHQSEBELSSQPGDF 1198  
QY 474 -----TKCCTESLVNRRPCFSALAEVDETY-----VPKEFNAETF--- 507  
Db 1199 SEEAEDSQCCS-----FKLLVEEGYEADSENPEDESETWDDGVDLKPFAESFIAS 1249  
QY 508 -----TFHADICTIS-----EKERQI----- 523

Db 1250 SSPNDLLENLSQGEIITYPEICTLELNLLSTGKAKLDVLAHVFESEFLKIIRQKEKIFLLM 1309  
QY 524 -----KQOTALVELVXHKPKA 539  
Db 1310 QOGTVKNLLGGFLSILTQDSDFOACQORVLVDLLVSLMSSRTCSBELTLRLRIFLEKSPC 1369  
QY 540 TKEQLKAVM-----DDFAAF-----VEKC----- 558  
Db 1370 TEILLGLIKLIVESDITMSPSQYLTFFLLHTPNLSNGVSSQKPCGILNSKAMGLLRARV 1429  
QY 559 -----CKADDKETCFAE 570  
Db 1430 SOSKIEGSESEFPQOOLSSWHIAPVHLPLLGOPHSEGSISLWNVNVECIHEPEST--TE 1487  
QY 571 EGKK-----LVAASQAALGLMSPRLEVPCHALPOG---LSPGOVIVRGLV----- 614  
Db 1488 KGKTRKRNKSLVLLDSSFDGTENNRLE-----GAAVYNPGERLIEGCVHMISLG 1538  
QY 615 -----LOEP 618  
Db 1539 SKALIQWADPHTGTFTIFRVCMDSDNDTKVLLAQVESQENIFLPSKWQHLVLTVLOQP 1598  
QY 619 K-----HFTVSLRDQAAHAPVTLRASFADRTLQWISRWGQKKLISAPFLFYQRFEEVLL 673  
Db 1599 QOKNINHGISI-----WIS--GQRK-----P-----DVTL 1622  
QY 674 LF-----QEGGLKLA-----LNGOGLGATSMNQQA 698  
Db 1623 DFMLPRKTSLSSDSNKTFCMIGHCLSSQSESLQLAGKWDLGNLLLPNGAKIG---SQEA 1678  
QY 699 L-----EQLREL-----RISG 709  
Db 1679 FYLYACGNHTSIMPKYKGPVNDYSKYINTEILOCEQIRELFMTSKDVIDGLLIESLV 1738  
QY 710 SVQLYC 715  
Db 1739 VYTTYC 1744

Search completed: October 17, 2005, 08:36:37

Job time : 53 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2005, 08:25:57 ; Search time 178 Seconds  
(without alignments)  
2065.578 Million cell updates/sec

Title: US-10-933-523-18

Perfect score: 3785

Sequence: 1 DAHKEVAHRFKDLGEENFK.....leqirelriagsvqlcyvhs 718

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.1

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : UniProt\_03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3103	82.0	609	1	ALBU_HUMAN
2	3079	81.3	609	2	Q68DN5
3	2942	77.7	600	1	ALBU_MACMU
4	2620	69.2	508	1	ALBU_FELCA
5	2614	69.1	584	2	Q7YSG3
6	2562	67.7	608	1	ALBU_CANFA
7	2504	66.2	608	2	Q95VB7
8	2475.9	65.4	607	1	ALBU_HORSE
9	2457	64.9	608	1	ALBU_RABIT
10	2450.9	64.8	607	1	ALBU_BOVIN
11	2432.9	64.3	607	1	ALBU_SHEEP
12	2426	64.1	608	1	ALBU_RAT
13	2416.8	63.9	605	1	ALBU_PIG
14	2409.8	63.7	607	2	Q68NH7
15	2387	63.1	608	2	Q6WDN9
16	2387	63.1	609	1	ALBU_MERUN
17	2379.9	62.9	583	2	Q6B3Z0
18	2378	62.8	608	1	ALBU_MOUSE
19	2374	62.7	608	2	Q8C7H3
20	2336	61.7	576	2	Q8C7C7
21	2045.8	54.1	417	2	Q85VG0
22	1950.7	51.5	396	2	Q81UK7
23	1557.9	41.2	615	1	ALBU_CHICK
24	1296.7	34.3	527	2	Q8JIA9
25	1256.7	33.2	609	1	FETA_PANTR
26	1252.7	33.1	609	1	FETA_HUMAN
27	1249.3	33.0	626	2	Q8UW05
28	1245.7	32.9	609	1	FETA_GORGO
29	1242.8	32.8	609	2	Q8MJU5
30	1219.7	32.2	610	2	Q8MJ76
31	1206.6	31.9	607	1	ALB2_XENLA

ALIGNMENTS

32	1201.2	31.7	609	1	FETA_HORSE
33	1200.6	31.7	607	2	Q642F7
34	1183.9	31.3	608	2	Q7TSE3
35	1168.1	30.9	606	1	ALB1_XENLA
36	1096.8	29.0	624	2	Q8UW06
37	1086.8	28.7	605	1	FETA_MOUSE
38	1084.8	28.7	605	2	Q8BK55
39	1078.8	28.5	605	2	Q8BK56
40	1071.6	28.3	611	1	FETA_RAT
41	1059	28.0	599	1	AFAM_HUMAN
42	1050.7	27.8	400	2	Q8JIA7
43	961.4	25.4	603	2	Q9YGH6
44	948.7	25.1	611	2	AFAM_MOUSE
45	933.4	24.7	614	2	Q91134
46	932	24.6	608	1	AFAM_RAT
47	894.6	23.6	406	2	Q8JIA8
48	771	20.4	167	2	Q6UXK4
49	765.3	20.2	608	1	ALB1_SALSA
50	761.3	20.1	608	2	ALB2_SALSA
51	739	19.5	205	2	Q8CG74
52	716.9	18.9	417	2	Q8R0J9
53	709.5	18.7	336	1	LEGC_HUMAN
54	704.9	18.6	296	2	Q9NZ03
55	700.4	18.5	382	1	ALBU_RANCA
56	601.7	15.9	314	1	LEGC_MOUSE
57	531.4	14.0	1423	1	ALBU_PETNA
58	527.7	13.9	383	2	Q8C3K0
59	421	11.1	476	1	VTDB_RAT
60	421	11.1	476	2	Q68FY4
61	417.5	11.0	474	1	VTDB_HUMAN
62	415.2	11.0	551	2	Q42279
63	414.8	11.0	476	1	VTDB_RABIT
64	414	10.9	476	1	VTDB_MOUSE
65	408.5	10.8	484	2	Q9W6F5
66	375.3	9.9	3207	2	Q8MW03
67	359.7	9.5	3287	2	Q8T326
68	347.4	9.2	3542	2	Q9U5M2
69	344.9	9.1	3279	2	Q9N4B9
70	343.9	9.1	3843	2	Q9VU94
71	339.8	9.0	464	2	Q6DGV8
72	338.9	9.0	3843	2	Q9U5D0
73	335	8.9	1797	2	Q75X61
74	332.2	8.8	2002	2	Q75XH1
75	332.1	8.8	3796	2	Q97699
76	331.6	8.8	3666	2	Q6UDX0
77	331.5	8.8	3796	2	Q9TTK4
78	331.1	8.7	4283	2	Q8UWL7
79	331	8.7	122	2	Q90WZ8
80	329.6	8.7	1927	2	Q25262
81	329	8.7	1802	2	Q75X05
82	328.7	8.7	1820	2	Q75XM5
83	328.6	8.7	1813	2	Q75XE3
84	328.2	8.7	2316	2	Q8WXA6
85	327.2	8.6	1819	2	Q9ZLV0
86	326.9	8.6	2858	2	Q8IAK2
87	326.8	8.6	2443	2	Q96J17
88	326.6	8.6	2994	2	Q81I24
89	325	8.6	1927	2	Q75WX9
90	324.9	8.6	2118	1	ESP1_MOUSE
91	324.7	8.6	1894	2	Q75XB6
92	324.5	8.6	3954	2	Q6LEV2
93	324.3	8.6	1795	2	Q75XJ8
94	324	8.6	2154	2	Q9ZVC8
95	323.5	8.5	2342	2	Q8NG31
96	323.3	8.5	6885	1	SNE2_HUMAN
97	322.9	8.5	2553	2	Q6PUE8
98	322.9	8.5	3175	2	Q7QPT9
99	322.8	8.5	2418	1	SPCA_HUMAN
100	322.5	8.5	3346	2	Q8IEU9

Fused



RESULT 1  
ALBU\_HUMAN STANDARD; PRT; 609 AA.  
AC P02768; O95574; Q13140; Q9P157; Q9P117; Q9UHS3; Q9UJ20;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Serum albumin precursor.  
GN Name=ALB;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86196112; PubMed=3009475;  
RA Minghetti P.P., Ruffner D.E., Kuang W.J., Dennison O.E., Hawkins J.W.,  
RA Beattie W.G., Dugaiczak A.;  
RT "Molecular structure of the human albumin gene is revealed by  
RT nucleotide sequence within q11-22 of chromosome 4.";  
RL J. Biol. Chem. 261:6747-6757(1986).  
RN [2]  
RP SEQUENCE FROM N.A., AND VARIANT LYS-420.  
RX MEDLINE=82081882; PubMed=6171778;  
RA Lawn R.M., Adelman J., Bock S.C., Franke A.E., Houck C.M.,  
RA Najarian R.C., Seeburg P.H., Wion K.L.;  
RT "The sequence of human serum albumin cDNA and its expression in E.  
RT coli.";  
RL Nucleic Acids Res. 9:6103-6114(1981).  
RN [3]  
RP SEQUENCE FROM N.A., AND VARIANT GLY-121.  
RX MEDLINE=82105994; PubMed=6275391;  
RA Dugaiczak A., Law S.W., Dennison O.E.;  
RT "Nucleotide sequence and the encoded amino acids of human serum  
RT albumin mRNA.";  
RL Proc. Natl. Acad. Sci. U.S.A. 79:71-75(1982).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RA Yang S., Zhang R.A., Qi Z.W., Yuan Z.Y.;  
RL Submitted (SEP-1999) to the EMBL/GenBank/DBSJ databases.  
RN [5]  
RP SEQUENCE FROM N.A. (PRO0903/PRO1708/PRO2044/PRO2619/PRO2675).  
RC TISSUE=Fetal liver;  
RA Zhang C., Yu Y., Zhang S., Wei H., Bi J., Zhou G., Dong C., Zai Y.,  
RA Xu W., Gao F., Liu M., He F., Zhang Y., Ouyang S., Luo L.;  
RT "Functional prediction of the coding sequences of 121 new genes  
RT deduced by analysis of cDNA clones from human fetal liver.";  
RL Submitted (FEB-1999) to the EMBL/GenBank/DBSJ databases.  
RN [6]  
RP SEQUENCE FROM N.A., AND VARIANT HIROSHIMA-1 LYS-378.  
RA Huang M.C., Wu H.T.;  
RT "The cDNA sequences of human serum albumin.";  
RL Submitted (AUG-2002) to the EMBL/GenBank/DBSJ databases.  
RN [7]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver, and Skeletal muscle;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J.J., Helton E., Kettman M., Madan A.C., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A.C., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [8]  
RP SEQUENCE OF 25-609.  
RX MEDLINE=76187907; PubMed=1225573; DOI=10.1016/0014-5793(75)80242-0;  
RA Meloun B., Moravsek L., Kostka V.;  
RT "Complete amino acid sequence of human serum albumin.";  
RL FEBS Lett. 58:134-137(1975).  
RN [9]  
RP SEQUENCE OF 25-609.  
RA Brown J.R., Shockley P., Behrens P.Q.;  
RL (in) Bing D.H. (eds.);  
RL The chemistry and physiology of the human plasma proteins, pp.23-40,  
RL Pergamon Press, New York (1979).  
RN [10]  
RP SEQUENCE OF 1-455 FROM N.A.  
RC TISSUE=Liver;  
RA Menaya J., Farrilla R., Ayuso M.S.;  
RL Submitted (MAR-1995) to the EMBL/GenBank/DBSJ databases.  
RN [11]  
RP SEQUENCE OF 1-26 FROM N.A.  
RX MEDLINE=86140099; PubMed=2419329;  
RA Urano Y., Watanabe K., Sakai M., Tamaoki T.;  
RT "The human albumin gene. Characterization of the 5' and 3' flanking  
RT regions and the polymorphic gene transcripts.";  
RL J. Biol. Chem. 261:3244-3251(1986).  
RN [12]  
RP SEQUENCE OF 222-229.  
RX MEDLINE=76257808; PubMed=955075; DOI=10.1016/0014-5793(76)80496-6;  
RA Walker J.E.;  
RT "Lysine residue 199 of human serum albumin is modified by  
RT acetylsalicylic acid.";  
RL FEBS Lett. 66:173-175(1976).  
RN [13]  
RP SEQUENCE OF 25-44 AND 480-499.  
RC TISSUE=Heart;  
RX MEDLINE=95203287; PubMed=7895732;  
RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;  
RT "The human myocardial two-dimensional gel protein database: update  
RT 1994.";  
RL Electrophoresis 15:1459-1465(1994).  
RN [14]  
RP DISULFIDE BONDS.  
RA Sabar M.A., Stockbauer P., Moravsek L., Meloun B.;  
RT "Disulfide bonds in human serum albumin.";  
RL Collect. Czech. Chem. Commun. 42:564-579(1977).  
RN [15]  
RP BILIRUBIN-BINDING SITE.  
RX MEDLINE=78186630; PubMed=656055;  
RA Jacobsen C.;  
RT "Lysine residue 240 of human serum albumin is involved in high-  
RT affinity binding of bilirubin.";  
RL Biochem. J. 171:453-459(1978).  
RN [16]  
RP VARIANT CANTERBURY ASN-337.  
RX MEDLINE=87157744; PubMed=3828358; DOI=10.1016/0167-4838(87)90088-4;  
RA Brennan S.O., Herbert P.;  
RT "Albumin Canterbury (313 Lys-->Asn). A point mutation in the second  
RT domain of serum albumin.";  
RL Biochim. Biophys. Acta 912:191-197(1987).  
RN [17]  
RP VARIANTS NAG-2 AND NAG-3.  
RX MEDLINE=88068523; PubMed=3479777;  
RA Takahashi N., Takahashi Y., Isobe T., Putnam F.W., Fujita M.,  
RA Satoh C., Neel J.V.;  
RT "Amino acid substitutions in inherited albumin variants from  
RT Amerindian and Japanese populations.";  
RL Proc. Natl. Acad. Sci. U.S.A. 84:8001-8005(1987).  
RN [18]

RP VARIANTS NAG-1; HIR-1; HIR-2 AND TOCHIGI.  
 RX MEDLINE=89345611; PubMed=2762316;  
 RA Arai K., Madison J., Huse K., Ishioka N., Satoh C., Fujita M.,  
 RA Neel J.V., Sakurabayashi I., Putnam F.W.;  
 RT "Point substitutions in Japanese alloalbumins";  
 RL Proc. Natl. Acad. Sci. U.S.A. 86:6092-6096(1989).  
 [19]  
 RP VARIANTS MANAUS; OSARA; NAGOYA; FUKUOKA; HONOLULU AND NEW-GUINEA.  
 RX MEDLINE=90115905; PubMed=2404284;  
 RA Arai K., Madison J., Shimizu A., Putnam F.W.;  
 RT "Point substitutions in albumin genetic variants from Asia";  
 RL Proc. Natl. Acad. Sci. U.S.A. 87:497-501(1990).  
 [20]  
 RP DESCRIPTION OF VARIANT REDHILL.  
 RX MEDLINE=90115852; PubMed=2104980;  
 RA Brennan S.O., Myles T., Peach R.J., Donaldson D., George P.M.;  
 RT "Albumin Redhill (-1 Arg, 320 Ala-->Thr): a glycoprotein variant of  
 human serum albumin whose precursor has an aberrant signal peptidase  
 cleavage site";  
 RL Proc. Natl. Acad. Sci. U.S.A. 87:26-30(1990).  
 [21]  
 RP VARIANTS TORINO LYS-84; VARESE HIS-23 AND VIBO VALENTIA LYS-106.  
 RX MEDLINE=91062352; PubMed=2247440;  
 RA Galliano M., Minchiotti L., Porta F., Rossi A., Ferri G., Madison J.,  
 RA Watkins S., Minchiotti L., Putnam F.W.;  
 RT "Mutations in genetic variants of human serum albumin found in  
 Italy";  
 RL Proc. Natl. Acad. Sci. U.S.A. 87:8721-8725(1990).  
 [22]  
 RP VARIANT VENEZIA.  
 RX MEDLINE=91296740; PubMed=2068071;  
 RA Watkins S., Madison J., Davis E., Sakamoto Y., Galliano M.,  
 RA Minchiotti L., Putnam F.W.;  
 RT "A donor splice mutation and a single-base deletion produce two  
 carboxyl-terminal variants of human serum albumin";  
 RL Proc. Natl. Acad. Sci. U.S.A. 88:5959-5963(1991).  
 [23]  
 RP VARIANTS IOWA CITY-2 VAL-25; IOWA CITY-1 VAL-389; KOMAGOME-3 HIS-23;  
 RX MEDLINE=92052189; PubMed=1946412;  
 RA Madison J., Arai K., Feld R.D., Kyle R.A., Watkins S., Davis E.,  
 RA Matsuda Y.-I., Anaki I., Putnam F.W.;  
 RT "Genetic variants of serum albumin in Americans and Japanese";  
 RL Proc. Natl. Acad. Sci. U.S.A. 88:9853-9857(1991).  
 [24]  
 RP VARIANT CASEBOOK ASN-518.  
 RX MEDLINE=91316157; PubMed=1859851; DOI=10.1016/0925-4439(91)90023-3;  
 RA Peach R.J., Brennan S.O.;  
 RT "Structural characterization of a glycoprotein variant of human serum  
 albumin: albumin Casebrook (494 Asp-->Asn).";  
 RL Biochim. Biophys. Acta 1097:49-54(1991).  
 [25]  
 RP VARIANTS SONDRIO LYS-357 AND PARIS-2 ASN-587.  
 RX MEDLINE=92190239; PubMed=1347703; DOI=10.1016/0167-4838(92)90207-T;  
 RA Minchiotti L., Galliano M., Stoppini M., Ferri G., Crespeau H.,  
 RA Rochu D., Porta F.;  
 RT "Two alloalbumins with identical electrophoretic mobility are produced

181 KLDLRDEGKASSAKQRLKASLQKFGERAFKAWAVARLSORFPAEFAEYSLVDTLTK 240  
 205 KLDLRDEGKASSAKQRLKASLQKFGERAFKAWAVARLSORFPAEFAEYSLVDTLTK 264  
 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCPEKPLEKSHCIAEVENDEMPA 300  
 265 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCPEKPLEKSHCIAEVENDEMPA 324  
 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFYIYVARRHPDYSVLLRLAKTYETTLK 360  
 325 DLPSLAADFVESKDVCKNYAEAKDVFGLMFYIYVARRHPDYSVLLRLAKTYETTLK 384  
 361 CAAADPHCYAKVDFEKPFLVEEPQNLIKQNCLEFQEQGEYKFNALAVRYTKKVPQVST 420  
 385 CAAADPHCYAKVDFEKPFLVEEPQNLIKQNCLEFQEQGEYKFNALAVRYTKKVPQVST 444  
 421 PTLVEVSRLKGVSKCKCHPEAKRMPCAEDYLSVNLQNLVLEHKTPTVSDRVTKCCTES 480  
 445 PTLVEVSRLKGVSKCKCHPEAKRMPCAEDYLSVNLQNLVLEHKTPTVSDRVTKCCTES 504  
 481 LVNRRPCFSALAEVDYVYVPEFNAETFTFHADICTLSEKROIKQTALVELVKKHKPAT 540  
 505 LVNRRPCFSALAEVDYVYVPEFNAETFTFHADICTLSEKROIKQTALVELVKKHKPAT 564  
 541 KEQLKAVNMDFAAFVEKCKKADDKETCFPAEKGKLVAAASQAALGL 585  
 565 KEQLKAVNMDFAAFVEKCKKADDKETCFPAEKGKLVAAASQAALGL 609

RESULT 2  
 Q68DN5  
 ID Q68DN5 PRELIMINARY; PRT; 609 AA.  
 AC Q68DN5;  
 DT 25-OCT-2004 (TRENBLrel. 28, Created)  
 DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TRENBLrel. 28, Last annotation update)  
 DE Hypothetical protein DKFZp779N1935.  
 GN Name=DKFZp779N1935;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RG The German cDNA Consortium;  
 RA Bloecker H., Boecher M., Brandt P., Mewes H.W., Weil B., Amid C.,  
 RA Osanger A., Fobo G., Han M., Wiemann S.;  
 RL Submitted (AUG-2004) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; CR749331; CAH18185.1;  
 DR InterPro; IPR001703; Alphafetoprot.  
 DR InterPro; IPR000264; Serum albumin.  
 DR Pfam; PF00273; Serum albumin; 3.  
 DR PRINTS; PRO0803; AFETOPROTEIN.  
 DR PRINTS; PRO0802; SERUMALBUMIN.  
 DR ProDom; PD002486; Serum albumin; 1.  
 DR SMART; SM00103; ALBUMIN; 3.  
 DR PROSITE; PS00212; ALBUMIN; 3.  
 KW Hypothetical protein.  
 SQ SEQUENCE 609 AA; 69402 MW; 3BA3AFF17BF99E94 CRC64;

Query Match 81.3%; Score 3079; DB 2; Length 609;  
 Best Local Similarity 99.1%; Pred. No. 1.2e-69;  
 Matches 580; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

1 DAHSEVAHFRKDLGEENFKALVLIAPQYLOQCPEFHVKLVNEVTEFAKTCVADESAE 60  
 25 DAHSEVAHFRKDLGEENFKALVLIAPQYLOQCPEFHVKLVNEVTEFAKTCVADESAE 84  
 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKKDNDPNLRLVRPEV 120  
 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKKDNDPNLRLVRPEV 144  
 121 DVNCTAFHNEETFLKKLYEYIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180  
 145 DVNCTAFHNEETFLKKLYEYIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 204

```
Qy 121 DVMCTAFHDNEETFLKKYLVEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP 180
Db 145 DVMCTAFHDNEETFLKKYLVEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP 204
Qy 181 KLDELDEGKASSAKQRLKASLOKFGERAFAKAWARLSORFPKABFAEVSKLVTDLTK 240
Db 205 KLDELDEGKASSAKQRLKASLOKFGERAFAKAWARLSORFPKABFAEVSKLVTDLTK 264
Qy 241 VHTCCGHDLLLECADRADLAKYICENQDISISSKLKCECKPLLEKSHCIAEVENDEMPA 300
Db 265 VHTCCGHDLLLECADRADLAKYICENQDISISSKLKCECKPLLEKSHCIAEVENDEMPA 324
Qy 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYEYARRHPDYSVLLLRLLAKTYETTLK 360
Db 325 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYEYARRHPDYSVLLLRLLAKTYETTLK 384
Qy 361 CAAADPHECAKVFDEPKPLVEEPQNLIKONCELFEQLGEYKFQNALLVRYTKVPQVST 420
Db 385 CAAADPHECAKVFDEPKPLVEEPQNLIKONCELFEQLGEYKFQNALLVRYTKVPQVST 444
Qy 421 PTLVEVRNLGKVGSKCKKHPEAKRMPCAEADYLSVNLNQLCVLHEKTPVSDRVTKCCTES 480
Db 445 PTLVEVRNLGKVGSKCKKHPEAKRMPCAEADYLSVNLNQLCVLHERTPVSDRVTKCCTES 504
Qy 481 LVNRRPCFSALEVDVETVPKFNATETTFHADICTLSEKEROIKKOTALVELVKGHPKAT 540
Db 505 LVNRRPCFSALEVDVETVPKFNATETTFHADICTLSEKEROIKKOTALVELVKGHPKAT 564
Qy 541 KEQLKAVNDDFAAFVEKCCAKDDKTCFAEKGKLVAAASQAALG 585
Db 565 KEQLKAVNDDFAAFVEKCCAKDDKTCFAEKGKLVAAASQAALG 609

RESULT 3
ALBU_MACMU
ID _ALBU_MACMU STANDARD; PRT; 600 AA.
AC Q28522;
DT 01-NOV-1997 (Rel. 35, Created)
DR 01-NOV-1997 (Rel. 35, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Serum albumin precursor (Fragment).
GN Name=ALB;
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheciae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93211971; PubMed=8460152;
RA Watkins S.A., Sakamoto Y., Madison J.M., Davis E.M., Smith D.G.,
RA Dwulet J., Putnam F.W.;
RT "cDNA and protein sequence of polymorphic macaque albumins that differ
in bilirubin binding";
RL Proc. Natl. Acad. Sci. U.S.A. 90:2409-2413 (1993).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; M90463; AAA36906.1; -.
DR PIR; A47391; A47391.
DR HSP; P02768; IE7B.
DR InterPro; IPR001703; Alphafetoprot.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; Serum_albumin; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Copper; Lipid-binding; Metal-binding; Repeat; Signal.
FT NON TER 1
FT SIGNAL <1 10 By similarity.
FT PROPEP 11 16 By similarity.
FT CHAIN 17 600 Serum albumin.
FT DOMAIN 17 197 Albumin 1.
FT DOMAIN 204 389 Albumin 2.
FT DOMAIN 396 587 Albumin 3.
FT METAL 19 19 Copper (By similarity).
FT BINDING 256 256 Bilirubin (Potential).
FT DISULFID 69 78 By similarity.
FT DISULFID 91 107 By similarity.
FT DISULFID 106 117 By similarity.
FT DISULFID 140 185 By similarity.
FT DISULFID 184 193 By similarity.
FT DISULFID 216 262 By similarity.
FT DISULFID 261 269 By similarity.
FT DISULFID 281 295 By similarity.
FT DISULFID 294 305 By similarity.
FT DISULFID 332 377 By similarity.
FT DISULFID 376 385 By similarity.
FT DISULFID 408 454 By similarity.
FT DISULFID 453 464 By similarity.
FT DISULFID 477 493 By similarity.
FT DISULFID 492 503 By similarity.
FT DISULFID 530 575 By similarity.
FT DISULFID 574 583 By similarity.
SQ SEQUENCE 600 AA; 67880 MW; E45C871A670E740B CRC64;

Query Match 77.7%; Score 2942; DB 1; Length 600;
Best Local Similarity 93.5%; Pred. No. 3.5e-66;
Matches 545; Conservative 23; Mismatches 15; Indels 0; Gaps 0;

Qy 1 DAHKSEVAHRPKDLGEENFKALVLIAPAOYLQCCPREDHVKLVNVEVTEFAKTCVADSSAE 60
Db 17 DTHKSEVAHRPKDLGEBEHFGLVLPFSQYLQCCPFEHVKLVNVEVTEFAKTCVADSSAE 76
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPVLVRPEV 120
Db 77 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPVLVRPEV 136
Qy 121 DVMTAFHDNEETFLKKYLVEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP 180
Db 137 DVMTAFHDNEETFLKKYLVEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP 196
Qy 181 KLDELDEGKASSAKQRLKASLOKFGERAFAKAWARLSORFPKABFAEVSKLVTDLTK 240
Db 197 KLDELDEGKASSAKQRLKASLOKFGERAFAKAWARLSORFPKABFAEVSKLVTDLTK 256
Qy 241 VHTCCGHDLLLECADRADLAKYICENQDISISSKLKCECKPLLEKSHCIAEVENDEMPA 300
Db 257 VHTCCGHDLLLECADRADLAKYICENQDISISSKLKCECKPLLEKSHCIAEVENDEMPA 316
Qy 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYEYARRHPDYSVLLLRLLAKTYETTLK 360
Db 317 DLPSLAADYVESKDVCKNYAEAKDVFGLMFLYEYARRHPDYSVLLLRLLAKTYETTLK 376
Qy 361 CAAADPHECAKVFDEPKPLVEEPQNLIKONCELFEQLGEYKFQNALLVRYTKVPQVST 420
Db 377 CAAADPHECAKVFDEPKPLVEEPQNLIKONCELFEQLGEYKFQNALLVRYTKVPQVST 436
Qy 421 PTLVEVRNLGKVGSKCKKHPEAKRMPCAEADYLSVNLNQLCVLHEKTPVSDRVTKCCTES 480
```

```
Db 437 PTLVEVSRLGKVGKACCKLPEAKMPCAEYLSVVLNRLCVLHKEKTPVSKVTKCTES 496
Qy 481 LVNRRPCFSALEVDYTPKFNATFTFHADICTLSEKEROIKKOTATLVELVHKPKAT 540
Db 497 LVNRRPCFSALELDRAVVPKAFNATFTFHADICTLSEKEROIKKOTATLVELVHKPKAT 556
Qy 541 KEQLKAVNMDFAAFVEKCKCADDKTCFAEGKKLVAAASQAAL 583
Db 557 KEQLKGVMDNFAAFVEKCKCADDKTCFAEGKPFVAAASQAAL 599

RESULT 4
ALBU_FELCA STANDARD; PRT; 608 AA.
AC P49064;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Serum albumin precursor (Allergen Fel d 2).
GN Name=ALB;
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96194824; PubMed=8647469; DOI=10.1016/0378-1119(95)00851-9;
RA Hilger C., Grigioni F., Kohnen M., Hentges F.;
RT "Sequence of the gene encoding cat (Felis domesticus) serum albumin.";
RL Gene 169:295-296(1996).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- ALLERGEN: Causes an allergic reaction in human.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; X84842; CAA59279.1; -
CC PIR; JC4660; S57632.
CC
CC HSSP; P02768; 1E7B.
CC
CC InterPro; IPR000264; Serum albumin.
CC Pfam; PF00273; Serum albumin; 3.
CC PRINTS; P00802; SERUMALBUMIN.
CC ProDom; PD002486; Serum albumin; 1.
CC SMART; SM00103; ALBUMIN; 3.
CC
CC PROSITE; PS00212; ALBUMIN; 3.
CC Allergen; Copper; Lipid-binding; Metal-binding; Repeat; Signal.
KW SIGNAL
FT 1 18 By similarity.
FT PROPEP 19 24 By similarity.
FT CHAIN 25 608 Serum albumin.
FT DOMAIN 25 205 Albumin 1.
FT DOMAIN 212 397 Albumin 2.
FT DOMAIN 404 595 Albumin 3.
FT METAL 27 27 Copper.
FT DISULFID 77 86 By similarity.
FT DISULFID 99 115 By similarity.
FT DISULFID 114 125 By similarity.
FT DISULFID 148 193 By similarity.
FT DISULFID 192 201 By similarity.
FT DISULFID 224 270 By similarity.
FT DISULFID 269 277 By similarity.
FT DISULFID 289 303 By similarity.
```

```
FT DISULFID 302 313 By similarity.
FT DISULFID 340 385 By similarity.
FT DISULFID 384 393 By similarity.
FT DISULFID 416 462 By similarity.
FT DISULFID 461 472 By similarity.
FT DISULFID 485 501 By similarity.
FT DISULFID 500 511 By similarity.
FT DISULFID 538 583 By similarity.
FT DISULFID 582 591 By similarity.
SQ SEQUENCE 608 AA; 68659 MW; 07E629CAC5F605F CRC64;

Query Match 69.2%; Score 2620; DB 1; Length 608;
Best Local Similarity 82.0%; Pred. No. 5.8e-58;
Matches 478; Conservative 52; Mismatches 53; Indels 0; Gaps 0;

Qy 1 DAHSEVAHREFKDLGEENFKALVLIAPAYLQQCPFDHVKLVNEVTEFAKTCVADESAB 60
Db 25 EAHQSEIAHRNFDLGEHFRGLVAVFSQYLQQCPFDHVKLVNEVTEFAKGCVAQSA 84
Qy 61 NCDKSLHTLFGDKLCTVATLTRETGYGEMADCCAKGPERNECFLOHKDDNPRLVRREV 120
Db 85 NCEKSLHELLGDKLCTVASLDRKYGEMADCCCKEPEERNECFLOHKDDNPFGQLVPEA 144
Qy 121 DVMCTAHNDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCOAAADKAACLLP 180
Db 145 DAMCTAFHENRQRFGLKYLVEIARRHPYFYAPELDYIABEYKGVFTECEAAADKAACLLP 204
Qy 181 KLDELDRSGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQRFKAEVSKLVTDLT 240
Db 205 KVDALREKVLASSAKERLKCSLQKFGGERAFKAWAVARLSQRFKAEVSKLVTDLTAK 264
Qy 241 VHTCCHGDLLECAADDRADLAKYICENQDSISSKKECKPPLLEKSHCIAEVNDMPA 300
Db 265 IHKECCHGDLLECAADDRADLAKYICENQDSISSTKLKECCGKPVLEKSHCISEVERDELPA 324
Qy 301 DLPSLAADPVESKOVKNYAEAKOVFLGCMFLYEYARRHPDYSVLLRLAKTYETTLEK 360
Db 325 DLPLAVDFVEDKEVKCNQYQAKOVFLGTFLEYYSRRHPPEYSVLLRLAKTYETTLEK 384
Qy 361 CAAADPCHCYAKVDFEFKPLVEEPONLIKQNCLEFEQLGEYKFNQNALVRYTKVPQVST 420
Db 385 CATDPPACYAHVDFEFKPLVEEPHNLVKTNCLEFEKLGEGFNQALLVRYTKVPQVST 444
Qy 421 PTLVEVSRLGKVGSKCKCHPEAKMPCAEYLSVVLNQLCVLHKEKTPVSDRVTKCTES 480
Db 445 PTLVEVSRLGKVGSKCKCTHPEAERLSCAEDYLSVVLNRLCVLHKEKTPVSDRVTKCTES 504
Qy 481 LVNRRPCFSALEVDYTPKFNATFTFHADICTLSEKEROIKKOTATLVELVHKPKAT 540
Db 505 LVNRRPCFSALEVDYTPKFNATFTFHADICTLSEKEROIKKOTATLVELVHKPKAT 564
Qy 541 KEQLKAVNMDFAAFVEKCKCADDKTCFAEGKKLVAAASQAAL 583
Db 585 KEQLKGVMDNFAAFVEKCKCADDKTCFAEGKPFVAAASQAAL 607

RESULT 5
QYSG3
ID QYSG3 PRELIMINARY; PRT; 584 AA.
AC QYSG3,
DT 01-OCT-2003 (TremBLrel. 25, Created)
DT 01-OCT-2003 (TremBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)
DE Albumin (Fragment).
GN Name=alb;
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Reininger R., Swoboda I., Bohle B., Hauswirth A.W., Valent P.,
```

RA Rumpold H., Valenta R., Spitzauer S.;  
 RL Submitted (WAY-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ487677; CAD322275.1; -.  
 DR HSP; P02768; 1E7B.  
 DR GO; GO:0005615; C:extracellular space; IEA.  
 DR GO; GO:0005386; F:carrier activity; IEA.  
 DR GO; GO:0006810; P:transport; IEA.  
 DR InterPro; IPR00264; Serum albumin.  
 DR Pfam; PF00273; Serum albumin.  
 DR ProDom; PD002486; Serum albumin; 1.  
 DR SMART; SM00103; ALBUMIN; 3.  
 DR PROSITE; PS00212; ALBUMIN; 3.  
 FT NON TER 1  
 SQ SEQUENCE 584 AA; 65908 MW; B51002F12902C9CE CRC64;  
 Query Match 69.1%; Score 2614; DB 2; Length 584;  
 Best Local Similarity 81.6%; Pred. No. 7.3e-53;  
 Matches 476; Conservative 54; Mismatches 55; Indels 0; Gaps 0;  
 Qy 1 DAHSEVAHPEKDLGEENFKALVIAFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESA 60  
 Db 1 EAHQSEIAHRNDLGEHFRGLVAVFSQYLQCCPFEDHVKLVNEVTEFAFGCVADQSA 60  
 Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVPEV 120  
 Db 61 NCEKSLHELFGDKLCTVASLDRKYGEMADCEKKEPERNECFLOHKDDNPFQGLVTP 120  
 Qy 121 DVMCTAFHNEETFLKKYLYEIARRHPYFYAPPELLFPAKRYKAAFTCCQAAADKAAC 180  
 Db 121 DAMCTAFHNEBQFLGKLYLYEIARRHPYFYAPPELLYAEYRGVFTCCCEAAADKA 180  
 Qy 181 KLDELREGKASSAKORLKCASLOKFGERAFAKAVARLSQRPKAPFAEVSKLVTDLTK 240  
 Db 181 KVDALREKVLASSAKERLKCASLOKFGERAFAKAVARLSQRPKAPFAEISKLVTDL 240  
 Qy 241 VHTCCCHGDLLECCADRDALAKYTCENODSISSKLKECCERKPLEKSHCIAEVENDEMP 300  
 Db 241 IHKECCCHGDLLECCADRDALAKYTCENODSISSKLKECCERKPLEKSHCISEVERDEL 300  
 Qy 301 DLPLSADFVSKOVCKNRYARAKVFLGMFLYEVARRHPDYSVLLRLAKTYTTTLEKC 360  
 Db 301 DLPLSADFVSKOVCKNRYARAKVFLGMFLYEVARRHPDYSVLLRLAKTYTTTLEKC 360  
 Qy 361 CAADPHCYAKVDFEKLPEEPQNLTKONCELFQELGEYKFNALLVRYTKVQPVST 420  
 Db 361 CATDPPACYAHVDFEKLPEEPQNLTKONCELFQELGEYKFNALLVRYTKVQPVST 420  
 Qy 421 PTLVEVSRNLGKVGSKCKHPKAPKMPCAEDYLSVNLNQLCVLHEKTPVSDRVTKCCTES 480  
 Db 421 PTLVEVSRNLGKVGSKCKHPKAPKMPCAEDYLSVNLNQLCVLHEKTPVSDRVTKCCTES 480  
 Qy 481 LVNRRPCFSALEVDYVYKPFNAETFTFHADICTLSEKQRIKKQTALVELVGHKPKAT 540  
 Db 481 LVNRRPCFSALEVDYVYKPFNAETFTFHADICTLSEKQRIKKQTALVELVGHKPKAT 540  
 Qy 541 KEQLKAVMDFAAFVEKCKKADDTETCAEESKGLVAASOAL 583  
 Db 541 BEQLKTVMGDFGFSVDKCAAEDEKAPFAEESKGLVAASOAL 583  
 RESULT 6  
 ID ALBU\_CANFA STANDARD; PRT; 608 AA.  
 AC P49822; O77705; Q9TS24;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Serum albumin precursor (Allergen Can f 3).  
 GN Name=ALB;  
 OS Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 OX NCBI\_TaxID=9615;

[1] SEQUENCE FROM N.A.  
 RC STRAIN=Beagle; TISSUE=Liver;  
 RA Hilger C.; Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
 [2]  
 RN SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=20148667; PubMed=10669848; DOI=10.1016/S0091-6749(00)90077-0;  
 RA Pandjaitan B., Swoboda I., Brandejsky-Pichler F., Rumpold H.,  
 Valenta R., Spitzauer S.;  
 RT "Escherichia coli expression and purification of recombinant dog  
 albumin, a cross-reactive animal allergen.";  
 RL J. Allergy Clin. Immunol. 105:279-285(2000).  
 [3]  
 RN SEQUENCE OF 25-48.  
 RX MEDLINE=75011422; PubMed=4414013;  
 RA Dixon J.W., Sarkar B.;  
 RT "Isolation, amino acid sequence and copper(II)-binding properties of  
 peptide (1-24) of dog serum albumin.";  
 RL J. Biol. Chem. 249:5872-5877(1974).  
 [4]  
 RN SEQUENCE OF 25-38.  
 RC TISSUE=Heart;  
 RX MEDLINE=98163340; PubMed=9504812;  
 RA Dunn M.J., Corbett J.M., Wheeler C.H.;  
 RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of  
 dog heart proteins.";  
 RL Electrophoresis 18:2795-2802(1997).  
 [5]  
 RN SEQUENCE OF 215-478 FROM N.A.  
 RC TISSUE=Salivary gland;  
 RX MEDLINE=94201492; PubMed=7512102;  
 RA Spitzauer S., Schweiger C., Speer W.R., Pandjaitan B., Valent P.,  
 Muehl S., Ebner C., Scheiner O., Kraft D., Rumpold H.;  
 RT "Molecular characterization of dog albumin as a cross-reactive  
 allergen.";  
 RL J. Allergy Clin. Immunol. 93:614-627(1994).  
 CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good  
 binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
 hormones, bilirubin and drugs. Its main function is the regulation  
 of the colloidal osmotic pressure of blood.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Plasma.  
 CC -!- ALLERGEN: Causes an allergic reaction in human.  
 CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.  
 CC -!- SIMILARITY: Contains 3 albumin domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 or send an email to license@sib-sib.ch).  
 EMBL; AJ133489; CAB64867.1; -.  
 EMBL; Y17737; CAAT8841.1; -.  
 EMBL; S72946; AAB30434.1; -.  
 HSP; P02768; 1E7B.  
 HSC-2DPAGE; P49822; DOG.  
 InterPro; IPR000264; Serum albumin.  
 Pfam; PF00273; Serum albumin; 3.  
 PRINTS; PR00802; SERUMALBUMIN.  
 ProDom; PD002486; Serum albumin; 1.  
 SMART; SM00103; ALBUMIN; 3.  
 PROSITE; PS00212; ALBUMIN; 3.  
 Allergen; Copper; Direct protein sequencing; Lipid-binding;  
 Metal-binding; Repeat; Signal.  
 SIGNAL 1 18 Potential.  
 PROPEP 19 24 Serum albumin.  
 CHAIN 25 608  
 DOMAIN 25 205 Albumin 1.

```
FT DOMAIN 212 397 Albumin 2.
FT DOMAIN 404 595 Albumin 3.
FT METAL 27 27 Copper (By similarity).
FT DISULFID 77 86 By similarity.
FT DISULFID 99 115 By similarity.
FT DISULFID 114 125 By similarity.
FT DISULFID 148 193 By similarity.
FT DISULFID 192 201 By similarity.
FT DISULFID 224 270 By similarity.
FT DISULFID 269 277 By similarity.
FT DISULFID 289 303 By similarity.
FT DISULFID 302 313 By similarity.
FT DISULFID 340 385 By similarity.
FT DISULFID 384 393 By similarity.
FT DISULFID 416 462 By similarity.
FT DISULFID 461 471 By similarity.
FT DISULFID 485 501 By similarity.
FT DISULFID 500 511 By similarity.
FT DISULFID 538 583 By similarity.
FT DISULFID 582 591 By similarity.
FT CONFLICT 1 26 MKWVFISLFLFSSAYSGRLVRREA -> MDT (in Ref. 2).
FT CONFLICT 146 146 A -> R (in Ref. 2).
FT CONFLICT 206 206 I -> T (in Ref. 2).
FT CONFLICT 349 349 V -> A (in Ref. 2).
FT CONFLICT 359 359 S -> A (in Ref. 1).
FT CONFLICT 448 448 V -> VV (in Ref. 5).
FT CONFLICT 474 474 D -> E (in Ref. 1).
SQ SEQUENCE 608 AA; 68606 MW; 3CF1C9FF7DD8FC06 CRC64;

Query Match 67.7%; Score 2562; DB 1; Length 608;
Best Local Similarity 79.8%; Pred. No. 1.7e-56;
Matches 465; Conservative 57; Mismatches 61; Indels 0; Gaps 0;

QY 1 DAHKEVAHRPKDGEENFKALVLAFAQYLOQCPFEHDHVLNVETSPAKTCVADESAG 60
DB 25 EAYKSEIARHYNDGEEHPRGLVAFASQYLOQCPFEHDHVLNVETSPAKTCVADESAG 84

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPPLRLVRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPPLRLVRPEV 144

QY 121 DVMCTAFHNDNEETFLKKLYEYIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLIP 180
DB 145 DALCAAFQDNEQLFLGKLYEYIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLIP 204

QY 181 KLDELDEGKASSAKQRLKCSLOKGERAFKAWAVARLSQRPKAEFAEVSCLVTDLTG 240
DB 205 KIEALREKVLSSAKERPKCASLOKGERAFKAWAVARLSQRPKAEFAEVSCLVTDLTG 264

QY 241 VHECCHGDLLECADRADLAKYICENODSISLKECEKPLEKSHCIAEVENDEMPA 300
DB 265 VHECCHGDLLECADRADLAKYICENODSISLKECEKPLEKSHCIAEVENDEMPA 324

QY 301 DLPSLAADFVESKQVCKNYAEAKDVFLGMFLYFYARRHPDYSVVLLRLAKTYETTTLEKC 360
DB 325 DLPSLAADFVEDKEVCKNYAEAKDVFLGMFLYFYARRHPDYSVVLLRLAKTYETTTLEKC 384

QY 361 CAADAPHECYAKVDFEFLVEEPQNLIKQNCLEFGEQYKFNALLVRYTKKQPVST 420
DB 385 CATDDPPTCYAKVDFEFLVEEPQNLIKQNCLEFGEQYKFNALLVRYTKKQPVST 444

QY 421 PTLVEVSRNLGKVGSKCKKHPRKMPCAEDVLSVVLNOLCVLHEKTPVSDRVTKCCTES 480
DB 445 PTLVEVSRNLGKVGSKCKKHPRKMPCAEDVLSVVLNOLCVLHEKTPVSDRVTKCCTES 504

QY 481 LVNRRPCFSALEVDETYVPKFNAAETFTFHADICTLSEKERQIKKQTALVELVGHKPKAT 540
DB 505 LVNRRPCFSALEVDETYVPKFNAAETFTFHADICTLSEKERQIKKQTALVELVGHKPKAT 564

QY 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKLVAAQAAL 583
DB 565 DEQLTKVMGDGFAFVEKCKKADDKETCFABEGKLVAAQAAL 607
```

```
RESULT 7
Q95VB7 PRELIMINARY; PRT; 608 AA.
AC Q95VB7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Albumin.
OS Schistosoma mansoni (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC Schistosomatidae; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6183;
RN [1]
RP SEQUENCE FROM N.A.
RA Osman A., Asahi H., Stadecker M.J., LoVerde P.T.;
RL Submitted (SRP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP418550; AAL08579.1; -.
DR HSRP; P02768; IHK1.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; P:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR001703; AlphaFoldProt.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; Serum_albumin; 3.
DR PRINTS; PR00803; AFETOPROTEIN.
DR PRINTS; PR00802; SERUMALBUMIN.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 608 AA; 68225 MW; E5EAB28E1C66E54 CRC64;

Query Match 66.2%; Score 2504; DB 2; Length 608;
Best Local Similarity 76.3%; Pred. No. 5.2e-55;
Matches 445; Conservative 79; Mismatches 59; Indels 0; Gaps 0;

QY 1 DAHKEVAHRPKDGEENFKALVLAFAQYLOQCPFEHDHVLNVETSPAKTCVADESAG 60
DB 25 DAHKEVAHRPKDGEENFKALVLAFAQYLOQCPFEHDHVLNVETSPAKTCVADESAG 84

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPPLRLVRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPPLRLVRPEV 144

QY 121 DVMCTAFHNDNEETFLKKLYEYIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLIP 180
DB 145 EAMCTSFQENAVTFMGHLYEVARHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLIP 204

QY 181 KLDELDEGKASSAKQRLKCSLOKGERAFKAWAVARLSQRPKAEFAEVSCLVTDLTG 240
DB 205 KLDELDEGKASSAKQRLKCSLOKGERAFKAWAVARLSQRPKAEFAEVSCLVTDLTG 264

QY 241 VHECCHGDLLECADRADLAKYICENODSISLKECEKPLEKSHCIAEVENDEMPA 300
DB 265 VHECCHGDLLECADRADLAKYICENODSISLKECEKPLEKSHCIAEVENDEMPA 324

QY 301 DLPSLAADFVESKQVCKNYAEAKDVFLGMFLYFYARRHPDYSVVLLRLAKTYETTTLEKC 360
DB 325 DLPSLAADFVEDKEVCKNYAEAKDVFLGMFLYFYARRHPDYSVVLLRLAKTYETTTLEKC 384

QY 361 CAADAPHECYAKVDFEFLVEEPQNLIKQNCLEFGEQYKFNALLVRYTKKQPVST 420
DB 385 CAADAPHECYAKVDFEFLVEEPQNLIKQNCLEFGEQYKFNALLVRYTKKQPVST 444

QY 421 PTLVEVSRNLGKVGSKCKKHPRKMPCAEDVLSVVLNOLCVLHEKTPVSDRVTKCCTES 480
DB 445 PTLVEVSRNLGKVGSKCKKHPRKMPCAEDVLSVVLNOLCVLHEKTPVSDRVTKCCTES 504

QY 481 LVNRRPCFSALEVDETYVPKFNAAETFTFHADICTLSEKERQIKKQTALVELVGHKPKAT 540
DB 505 LVNRRPCFSALEVDETYVPKFNAAETFTFHADICTLSEKERQIKKQTALVELVGHKPKAT 564

QY 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKLVAAQAAL 583
```







```

RN REVISIONS TO 322-323 AND 506-507.
RA Sheffield W.P.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by, and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U18344; AAB58347.2; -.
CC HSPF; P02768; 1E7B.
CC InterPro; IPR000264; Serum_albumin.
CC Pfam; PF00273; Serum_albumin; 3.
CC PRINTS; PR00802; SERUMALBUMIN.
CC ProDom; PD002486; Serum_albumin; 1.
CC SMART; SM00103; ALBUMIN_3.
CC PROSITE; PS00212; ALBUMIN; 3.
KW Copper; Lipid-binding; Metal-binding; Repeat; Signal.
FT SIGNAL 1 18 By similarity.
FT PROPEP 19 24 By similarity.
FT CHAIN 25 608 Serum albumin.
FT DOMAIN 25 205 Albumin 1.
FT DOMAIN 212 397 Albumin 2.
FT DOMAIN 404 595 Albumin 3.
FT METAL 27 27 Copper.
FT DISULFID 77 86 By similarity.
FT DISULFID 99 115 By similarity.
FT DISULFID 114 125 By similarity.
FT DISULFID 148 193 By similarity.
FT DISULFID 192 201 By similarity.
FT DISULFID 224 270 By similarity.
FT DISULFID 269 277 By similarity.
FT DISULFID 289 303 By similarity.
FT DISULFID 302 313 By similarity.
FT DISULFID 340 385 By similarity.
FT DISULFID 384 393 By similarity.
FT DISULFID 416 462 By similarity.
FT DISULFID 461 472 By similarity.
FT DISULFID 485 501 By similarity.
FT DISULFID 500 511 By similarity.
FT DISULFID 538 583 By similarity.
FT DISULFID 582 591 By similarity.
SQ SEQUENCE 608 AA; 68909 MW; 9ECAFDA86B1EF09 CRC64;

Query Match 64.9%; Score 2457; DB 1; Length 608;
Best Local Similarity 74.3%; Pred. No. 8.2e-54;
Matches 434; Conservative 78; Mismatches 72; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFDKLGEEFNKALVIAFAQYLOQCPEFDHVKLVNEVTEFAKTCVADESAAE 60
Db 25 EAHKSEIAHRFNDVGEHEHIGLVITFSYQLQKCPVEEHAHLVKEVITDLAKACVADESAA 84

QY 61 NCDKSLHTLFGDKLCTVATRLRTYEMADCCAKQEPERNECFLOHKDDNPRLVRLVRPE 120
Db 85 NCDKSLHDFDGDICALPCLSDTYGDVADCCKEPERNECFLOHKDDNPRLVRLVRPE 144

QY 121 DVMCTAFHNEETFLKKLYEYARRHPYFAPPELLFFAKRYKAAFTCCOARADKACLLP 180
Db 145 DVLCKAFHDDKAFGHLYEYARRHPYFAPPELLFYAOKYKAILTECCAAADKGCALTP 204

QY 181 KLDELRLDECKASSAKQLKCAASLQKFGERAFKAWAVARLSQRFPAEFAEVSKLVTDLTK 240

```

```

Db 205 KLDALGKSLISAQERLCASIOKFGDRAYKAWALVLSQRFKADFTDISKIVTDLTK 264
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCERPLEKSHCIAEVENDEMPA 300
Db 265 VHKECCCHGDLLECCADRADLAKYICENQDSISSKLKECCERPLEKSHCIAEVENDEMPA 324
QY 301 DLPGLADFVSKDVCNKYAEAKOVFLGMFLYEVARRHPDYVSVLLLLBLAKTYETTLK 360
Db 325 GLPAVAEEFEDKDVCKNYEAKDLFLGKFLYEVARRHPDYVSVLLLLBLGKAYEATLKK 384
QY 361 CAADPHCEYAKVDFEKFPLVEEENLQKCELFQELGGEYKFNQALLVRYTKKVPQVST 420
Db 385 CATDDPHACYAKVDFEKFPLVEEENLQKCELFQELGGEYKFNQALLVRYTKKVPQVST 444
QY 421 PTLVEVSRLNKGKSGKCKHPKAEKMPKAEADYLSVLLNQLCVLHKEKTPVSDRVTKCCTES 480
Db 445 PTLVEVSRLNKGKSGKCKHPKAEKMPKAEADYLSVLLNQLCVLHKEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALVETVYVPKFNATFTFHADICTLSEKQIKKQATLVELVHKHPKAT 540
Db 505 LVDRPCFSALVETVYVPKFNATFTFHADICTLSEKQIKKQATLVELVHKHPKAT 564
QY 541 KEQLKAVMDDFAAFEVKCKKADDKETCFAEKGLVAAASQAALG 584
Db 565 NDQLKTVVGEFTALLDKCCSAEDKEACFAVGGPKLVESKATLG 608

RESULT 10
ALBU_BOVIN
ID ALBU_BOVIN STANDARD; PRT; 607 AA.
AC P02769; O02787;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Serum albumin precursor (Allergen Bos d 6) (BSA).
GN Name=ALB;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Holowachuk E.W., Stoltenborg J.K., Reed R.G., Peters T. Jr.;
RT "Bovine serum albumin: cDNA sequence and expression.";
RL Submitted (AUG-1991) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A., AND VARIANT THR-214.
RA Barry T., Power S., Gannon F.;
RT "The bovine serum albumin mRNA.";
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA TISSUE=Liver;
RL MEDLINE=21195144; PubMed=11298124;
RA Hilger C., Grigioni F., De Beaufort C., Michel G., Freilinger J.,
RA Hentges F.;
RT "Differential binding of IgG and IgA antibodies to antigenic
RL determinants of bovine serum albumin.";
RN [4]
RP SEQUENCE FROM N.A., AND VARIANT THR-214.
RA Wu H.T., Huang M.C.;
RT "The complete cDNA sequence of bovine serum albumin.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 1-32.
RA MEDLINE=80024278; PubMed=488109;
RX McGilivray R.T.A., Chung D.W., Davie E.W.;
RT "Biosynthesis of bovine plasma proteins in a cell-free system. Amino-
RT terminal sequence of preproalbumin.";

```

RL Eur. J. Biochem. 98:477-485 (1979).  
 RN [6]  
 RX SEQUENCE OF 19-28.  
 RP MEDLINE=77134075; PubMed=843354;  
 RA Patterson J.E., Geller D.M.;  
 RT "Bovine microsomal albumin: amino terminal sequence of bovine  
 RL proalbumin.";  
 RN [7]  
 RP Biochem. Biophys. Res. Commun. 74:1220-1226 (1977).  
 RA SEQUENCE, AND REVISIONS TO 118-119 AND 180.  
 RX MEDLINE=91083649; PubMed=2260975;  
 RA Hirayama K., Akashi S., Furuya M., Fukuhara K.-I.;  
 RT "Rapid confirmation and revision of the primary structure of bovine  
 RL serum albumin by ESIMS and Frit-FAB LC/MS.";  
 RN [8]  
 RP Biochem. Biophys. Res. Commun. 173:639-646 (1990).  
 RA SEQUENCE OF 25-424 AND 429-607, AND VARIANT THR-214.  
 RX Brown J.R.;  
 RT "Structure of bovine serum albumin.";  
 RL [9]  
 RP Fed. Proc. 34:591-591 (1975).  
 RA REVISIONS TO 190-195.  
 RX Brown J.R.;  
 RT Submitted (APR-1975) to the PIR data bank.  
 RL [10]  
 RP SEQUENCE OF 25-64.  
 RX PubMed=2379503;  
 RA Strawich E., Glimcher M.J.;  
 RT "Tooth 'enamelins' identified mainly as serum proteins. Major  
 RL 'enamelin' is albumin.";  
 RN [11]  
 RP Eur. J. Biochem. 191:47-56 (1990).  
 RA SEQUENCE OF 25-41.  
 RX MEDLINE=88267456; PubMed=3389500;  
 RA Hsieh J.C., Lin F.P., Tam M.F.;  
 RT "Electroblotting onto glass-fiber filter from an analytical  
 RL isoelectrofocusing gel: a preparative method for isolating proteins  
 for N-terminal microsequencing.";  
 RN [12]  
 RP Anal. Biochem. 170:1-8 (1988).  
 RA SEQUENCE OF 163-172.  
 RX PubMed=2474609;  
 RA Carraway R.E., Cochran D.E., Boucher W., Mitra S.P.;  
 RT "Structures of histamine-releasing peptides formed by the action of  
 RL acid proteases on mammalian albumin(s).";  
 RN [13]  
 RP J. Immunol. 143:1680-1684 (1989).  
 RA SEQUENCE OF 402-433.  
 RX MEDLINE=8203364; PubMed=7283978;  
 RA Reed R.G., Putnam F.W., Peters T. Jr.;  
 RT "Sequence of residues 400-403 of bovine serum albumin.";  
 RL Biochem. J. 191:867-868 (1980).  
 RN [14]  
 RP SEQUENCE OF 437-451.  
 RA Vilbois F.;  
 RL Submitted (AUG-1998) to Swiss-Prot.  
 RN [15]  
 RP DISULFIDE BONDS.  
 RA Brown J.R.;  
 RT "Structure of serum albumin: disulfide bridges.";  
 RL Fed. Proc. 33:1389-1389 (1974).  
 CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good  
 CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
 CC hormones, bilirubin and drugs. Its main function is the regulation  
 CC of the colloidal osmotic pressure of blood.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- TISSUE SPECIFICITY: Plasma.  
 CC -1- ALLERGEN: Causes an allergic reaction in human.  
 CC -1- SIMILARITY: Belongs to the ALB/AFP/VDB family.  
 CC -1- SIMILARITY: Contains 3 albumin domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration -  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; MT3993; AAA51411.1; -;  
 DR EMBL; X58989; CAA41735.1; -;  
 DR EMBL; Y17769; CAA76847.1; -;  
 DR EMBL; AF542068; AAN17824.1; -;  
 DR PIR; A38885; ABBOS.  
 DR HSSP; P02768; IHK1.  
 DR InterPro; IPR001703; Alphaferoprot.  
 DR InterPro; IPR000264; Serum\_albumin.  
 DR Pfam; PF00273; Serum\_albumin; 3.  
 DR PRINTS; PR00802; SERUMALBUMIN.  
 DR ProDom; PD002486; Serum\_albumin; 1.  
 DR SMART; SM00103; ALBUMIN; 3.  
 DR PROSITE; PS00212; ALBUMIN; 3.  
 DR Allergen; Copper; Direct protein sequencing; Lipid-binding;  
 KW Metal-binding; Polymorphism; Repeat; Signal.  
 KW SIGNAL 1 18  
 FT PROPEP 19 24  
 FT CHAIN 25 607 Serum albumin.  
 FT DOMAIN 25 204 Albumin 1.  
 FT DOMAIN 211 396 Albumin 2.  
 FT DOMAIN 403 594 Albumin 3.  
 FT METAL 27 27 Copper (By similarity).  
 FT DISULFID 77 86  
 FT DISULFID 99 115  
 FT DISULFID 114 125  
 FT DISULFID 147 192  
 FT DISULFID 191 200  
 FT DISULFID 223 269  
 FT DISULFID 268 276  
 FT DISULFID 288 302  
 FT DISULFID 301 312  
 FT DISULFID 339 384  
 FT DISULFID 383 392  
 FT DISULFID 415 461  
 FT DISULFID 460 471  
 FT DISULFID 484 500  
 FT DISULFID 499 510  
 FT DISULFID 537 582  
 FT DISULFID 581 590  
 FT VARIANT 214 214 A -> T.  
 FT CONFLICT 58 58 Missing (in Ref. 10).  
 FT CONFLICT 302 302 C -> K (in Ref. 8).  
 FT CONFLICT 304 305 KP -> PC (in Ref. 8).  
 FT CONFLICT 324 324 N -> D (in Ref. 8).  
 FT CONFLICT 394 395 ST -> TS (in Ref. 8).  
 FT CONFLICT 437 437 K -> R (in Ref. 14).  
 FT CONFLICT 493 494 SE -> ES (in Ref. 8).  
 SQ SEQUENCE 607 AA; 69293 MW; 39167DFE768585D4 CRC64;  
 Query Match 64.8%; Score 2450.9; DB 1; Length 607;  
 Best local Similarity 75.8%; Pred. No. 1.2e-53;  
 Matches 442; Conservative 71; Mismatches 69; Indels 1; Gaps 1;  
 QY 1 DAUKSEVAHFRKDLGEENFKALVLIAPAOYLQCCPFDDHVKLVNEVTEFAKTCVADESAAE 60  
 DB 25 DTHKSEVAHFRKDLGEENFKALVLIAPAOYLQCCPFDDHVKLVNEVTEFAKTCVADESAAE 84  
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 120  
 DB 85 GCEKSLHTLFGDELCKVASLRETYGDMADCCQEPERNECFLOHKDDNPNLPLVRPEV 143  
 QY 121 DVMTAFHDNEETFLKYLVEIARRHYFYFAPELLPFAKRYKAFTCCQADKAACLLP 180  
 DB 144 NTLCDPEKADKKFWGKYLVEIARRHYFYFAPELLPFAKRYKAFTCCQADKAACLLP 203  
 QY 181 KLDELREGKASSAKORLKCASLQKFGERAFKAWAVARLSQRPFAKRYKAFTCCQADKAACLLP 240

204 KIETMREKVLASSARQRLRCASIQKGERALKAWSVARLSQKFPKABFVETVLVTLTK 263  
241 VHTCCCHGDLLEACADRADLAKYICENODSISSKLECECEKPLEKSHCIAEVENDEMPA 300  
264 VHKCCCHGDLLEACADRADLAKYICDNQDTISSKLEKCCDPLKXSHCIAEVEKDAIPE 323  
301 DLPSLAADFVESKDVCKNYABAKOVFLGMFLYEVARRHPDYVSVLLRLAKTYETTLK 360  
324 NLPLTADFAEDKDVCKNYQEAQDAFLGSLFYYSRRHPEYAVSVLLRLAKEYEATLEEC 383  
361 CAADPHCEYAKVDFEFPLVBEONLIKONCELFEOQLGEYKFNALLVRYTKVPOVST 420  
384 CAKDPHACYATVFDKLXHLVDEPNLIKONCELFKGEYGFONALIVRYTKVPOVST 443  
421 PTLVEVSRLGKVGSKCKKHPKAMPKCAEDYLSVNLQCLVLEKTPVSDRVTKCCTES 480  
444 PTLVEVSRLGKVGSKCKKHPKAMPKCAEDYLSVNLQCLVLEKTPVSEKVKCCTES 503  
481 LVNRRPCFSALEVDYVYKPEFNAFTFTFHADICTLSEKEROIKQTALVELVKKPKAT 540  
504 LVNRRPCFSALEVDYVYKPEFNAFTFTFHADICTLSEKEROIKQTALVELVKKPKAT 563  
541 KEOLKAVMDDDFAAFVEKCKCKADKTCFAEEGKLVAAASQAAL 583  
564 BEOLKTVMENFVAFVDDKCAADDKCAEFAVBPGLVSTQAL 606

## RESULT 11

ALBU SHEEP  
ID ALBU SHEEP STANDARD; PRT; 607 AA.  
AC P14639;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Serum albumin precursor.  
GN Name=ALB;  
OS Ovis aries (Sheep).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Caprinae; Ovis.  
OX NCBI\_TaxID=9940;  
RN  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=9009888; PubMed=2602160;  
RT Brown W.M., Driegielewska K.M., Foreman R.C., Saunders N.R.;  
RL "Nucleotide and deduced amino acid sequence of sheep serum albumin.";  
Nucleic Acids Res. 17:10495-10495(1989).  
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good  
binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
hormones, bilirubin and drugs. Its main function is the regulation  
of the colloidal osmotic pressure of blood.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Plasma.  
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.  
CC -!- SIMILARITY: Contains 3 albumin domains.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X17055; CAA34903.1; -.  
DR PIR; S06936; ABSHS.  
DR HSSP; P02768; LHK1.  
DR InterPro; IPR001703; Alphafetoprot.  
DR InterPro; IPR000264; Serum\_albumin.  
DR Pfam; PF00273; Serum\_albumin; 3.  
DR PRINTS; PR00802; SERUMALBUMIN.  
DR ProDom; PD002486; Serum\_albumin; 1.

Query Match 64.3%; Score 2432.9; DB 1; Length 607;  
Best Local Similarity 75.0%; Pred. No. 3.4e-53;  
Matches 437; Conservative 73; Mismatches 72; Indels 1; Gaps 1;  
QY 1 DAHSEVAHREFKDLGEENFKALVIAFAQYLQCCFFEDHVKLVNEVTFEAKTCVADESAE 60  
DB 25 DTHKSEIAHREFKDLGEENFKALVIAFAQYLQCCFFEDHVKLVNEVTFEAKTCVADESAE 84  
QY 61 NDCSLHPLFGDLCTVATLRETTGEMADCCAKQEPERNECFLOHKKDNPRLVVRPEV 120  
DB 85 GCDKSLHPLFGDLCTVATLRETTGEMADCCAKQEPERNECFLOHKKDNPRLVVRPEV 143  
QY 121 DVMCTAFHNEETFLKYLVEIARHPYFYAPPELLFFAKRYKAAFTCCQAAEDKAACLLP 180  
DB 144 DTLCAEFKADKKFWGKLYEVARRHPYFYAPPELLFFAKRYKAAFTCCQAAEDKAACLLP 203  
QY 181 KLDELRLDEGKASSAKQRLKCSAQKFGGERAFKAMAVARLSQRPFAEFAEVSCLVDTLTK 240  
DB 204 KIDAMREKVLASSARQRLRCASIQKGERALKAWSVARLSQKFPKADFTDVTIKVTLTK 263  
QY 241 VHTCCCHGDLLEACADRADLAKYICENODSISSKLECECEKPLEKSHCIAEVENDEMPA 300  
DB 264 VHKCCCHGDLLEACADRADLAKYICDNQDTISSKLEKCCDPLKXSHCIAEVEKDAIPE 323  
QY 301 DLPSLAADFVESKDVCKNYABAKOVFLGMFLYEVARRHPDYVSVLLRLAKTYETTLK 360  
DB 324 NLPLTADFAEDKDVCKNYQEAQDAFLGSLFYYSRRHPEYAVSVLLRLAKEYEATLEEC 383  
QY 361 CAADPHCEYAKVDFEFPLVBEONLIKONCELFEOQLGEYKFNALLVRYTKVPOVST 420  
DB 384 CAKDPHACYATVFDKLXHLVDEPNLIKONCELFKGEYGFONALIVRYTKVPOVST 443  
QY 421 PTLVEVSRLGKVGSKCKKHPKAMPKCAEDYLSVNLQCLVLEKTPVSDRVTKCCTES 480  
DB 444 PTLVEVSRLGKVGSKCKKHPKAMPKCAEDYLSVNLQCLVLEKTPVSEKVKCCTES 503  
QY 481 LVNRRPCFSALEVDYVYKPEFNAFTFTFHADICTLSEKEROIKQTALVELVKKPKAT 540  
DB 504 LVNRRPCFSALEVDYVYKPEFNAFTFTFHADICTLSEKEROIKQTALVELVKKPKAT 563  
QY 541 KEOLKAVMDDDFAAFVEKCKCKADKTCFAEEGKLVAAASQAAL 583  
DB 564 BEOLKTVMENFVAFVDDKCAADDKCAEFAVBPGLVSTQAL 606





```
25-OCT-2004 (TrEMBLrel. 28, Created)
25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
Albumin.
OS Sus scrofa (pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
NCBI_TaxID=9823;
[1]
SEQUENCE FROM N.A.
RA Sun S., Deng J., Zhou Y., Lu J., Wu X.;
RT "Porcine serum albumin gene."
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY663543; AAT98610.1; -
DR InterPro: IPR000264; Serum albumin.
DR Pfam: PF00273; Serum_albumin; 3.
DR PRINTS: PRO0802; SERUMALBUMIN.
DR ProDom: PD002486; Serum albumin; 1.
DR SMART: SM00103; ALBUMIN; 3.
DR PROSITE: PS00212; ALBUMIN; 3.
DR SEQUENCE 607 AA; 69691 MW; 7B8DAL3543CA99D8 CRC64;

Query Match 63.7%; Score 2409.8; DB 2; Length 607;
Best Local Similarity 71.7%; Pred. No. 1.3e-52;
Matches 440; Conservative 70; Mismatches 72; Indels 32; Gaps 2;

Qy 1 DAHSEVAHREPKDIGEENFKALVLIAPQVLYQQCFPEHDHVKLVNEVTEFAKTCVADSEAE 60
Db 25 DTYKSEIAHREPKDIGEQYFKGLVLIAPSQHLQQCFPEBEHVKLVREVTVEFAKTCVADSEAE 84
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLRVPEV 120
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLRVPEV 143
Qy 121 DVMTCTAHFNDNEFLKYLVEIARRHPYFYAPPELLFAKRYKAAFTCCQAADKAAACLLP 180
Db 144 VALCADFOEDQKFWGKYLVEIARRHPYFYAPPELLYAIYKQVFSECCQAADKAAACLLP 203
Qy 181 KLDELDEGKASSAKORLKCSAQKQGERAFKAWAVARLSORFPKAFBFAEVSCLVTDLT 240
Db 204 KIEHLREKVLTSAAKORLKCSAQKQGERAFKAWAVARLSORFPKADFTFETSKIVTDLAK 263
Qy 241 VHTCECHGDLLECCADDRADLAKYICENQDTSISKLKCECCPKLLEKSHCIAEVENDEMPA 300
Db 264 VHKCECHGDLLECCADDRADLAKYICENQDTISTKLKCCDRLLEKSHCIAEAKDELPA 323
Qy 301 DLPISLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYVSVLLRLAKTYETTLEKC 360
Db 324 DLNPLEHDFVEDKEVCNKYAEAKDVFLGTFLYEVARRHPDYVSVLLRLAKIYEATLEDC 383
Qy 361 CAAADPHECYAKVDFEKPFLVEEPQNLIKONCELFEOQLGEYKFNALLIVRYTKKVPQVST 420
Db 384 CAKEDPPACATVDFKQPLVDEPKNLIKONCELFKELGEYGFONALLIVRYTKKVPQVST 443
Qy 421 PTLVEVSRNLGKVGSKCKCKHPEAKRMPCAEADYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 444 PTLVEVARKLGLVSGRCCKRPEERLSCAEDYLSVLNQLCVLHEKTPVSEKVTCKCTES 503
Qy 481 LVNRRPCFSALEVDVETVYVPKEFNAETFTFHADICTLSEKEROIKKQATLVELVKKPKAT 540
Db 504 LVNRRPCFSALTPTDETYPKPFVEGTFTFHADLTLPDEKQIKKQATLVELLKKHHPAT 563
Qy 541 KEQLKAVMDDFAAVFEKCKADDKETCFABEGKLVLAASQAALGLMSRLEVPVPCSHALPQ 600
Db 564 BEQLRTVLGNFAAFVQKCCNAPDHACFAVGPX-----598
Qy 601 GLSPQGVIIIRGLV 614
Db 599 -----VIEIRGIL 606

RESULT 15
Q6WDN9
```



QY 541 KEQLKAVNDDFAAFVEKCKADKDKTCFAEBGKLVAAASQAAL 583  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 565 BEQMTXVNGDFAAFLEKCKCDDADNKEACFTEDGPKLVAKQATL 607

## RESULT 16

ALBU MERUN  
ID ALBU MERUN STANDARD; PRT; 609 AA.  
AC Q35090;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Serum albumin precursor.  
Name=ALB;  
OS Meriones unguiculatus (Mongolian jird) (Mongolian gerbil).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Gerbillinae;  
OC Meriones.  
OX NCBI\_TaxID=10047;  
RN [1]  
SEQUENCE FROM N.A.  
RC STRAIN=MGS IDR; TISSUE=Liver;  
-RX MEDLINE=9811663; PubMed=9455485;  
RA Yoshida K., Seto-Onshima A., Sinozara H.;  
RT "Sequencing of cDNA encoding serum albumin and its extrahepatic  
synthesis in the Mongolian gerbil, Meriones unguiculatus.";  
RL DNA Res. 4:351-354(1997).  
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good  
binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
hormones, bilirubin and drugs. Its main function is the regulation  
of the colloidal osmotic pressure of blood.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Plasma.  
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.  
CC -!- SIMILARITY: Contains 3 albumin domains.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AB006197; BAA21765.1; -;  
CC FIP; JC5838; JC5838.  
CC HSP; P02768; IE7B.  
CC InterPro; IPR001703; Alphafetoprot.  
CC InterPro; IPR000264; Serum albumin.  
CC Pfam; PF00273; Serum albumin; 3.  
CC PRINTS; PR00802; SERUMALBUMIN.  
CC ProDom; PD002486; Serum albumin; 1.  
CC SMART; SM00103; ALBUMIN; 3.  
CC PROSITE; PS00212; ALBUMIN; 3.  
KW Copper; Lipid-binding; Metal-binding; Repeat; Signal.  
FT SIGNAL 1 18 By similarity.  
FT PROPEP 19 24 By similarity.  
FT CHAIN 25 609 Serum albumin.  
FT DOMAIN 25 206 Albumin 1.  
FT DOMAIN 213 398 Albumin 2.  
FT DOMAIN 405 596 Albumin 3.  
FT METAL 28 28 Copper.  
FT FT 78 87 By similarity.  
FT DISULFID 100 116 By similarity.  
FT DISULFID 115 126 By similarity.  
FT DISULFID 149 194 By similarity.  
FT DISULFID 193 202 By similarity.  
FT DISULFID 225 271 By similarity.  
FT DISULFID 270 278 By similarity.  
FT DISULFID 290 304 By similarity.  
FT DISULFID 303 314 By similarity.  
FT DISULFID 341 386 By similarity.  
FT DISULFID 385 394 By similarity.

FT DISULFID 417 463 By similarity.  
FT DISULFID 462 473 By similarity.  
FT DISULFID 486 502 By similarity.  
FT DISULFID 501 512 By similarity.  
FT DISULFID 539 584 By similarity.  
FT DISULFID 583 592 By similarity.  
SQ SEQUENCE 609 AA; 68940 MW; 9CA5F97F675F1A48 CRC64;  
Query Match 63.1%; Score 2387; DB 1; Length 609;  
Best Local Similarity 73.9%; Pred. No. 5e-52;  
Matches 430; Conservative 65; Mismatches 87; Indels 0; Gaps 0;  
QY 2 AHKSEVAHRFDLGEENFKALVLAFAQYLQOCPEPDHVKLVNVEFPAKTCVADESSEN 61  
Db 27 AHKSEIAHRYKDLGKPKGLVLYTFQYLQKCSYEEHVKLVRVETDVFASCAKXDESSEN 86  
QY 62 CDKSLHTLFGDKLCTVATLRETYGEMADCCAKQPERNECFLOHKDDNPRLVPRPVD 121  
Db 87 CDKSLHTLFGDKLCSLFPNGEKAEMADCCAKQPERNECFLOHKDDNPQLPFPKAEAPD 146  
QY 122 VMCTAFHDNEETFLKKLYEIAARRHPYFAPPELLFFAKRYKAAFTTECCOAAADKAACLLPK 181  
Db 147 AMCTAFQENAEAFMGHYLHEVARRHPYFYGPELLYADKYTAVLTTECCAADDKGACLTAK 206  
QY 192 LDELDEGKASSAKORLKASLOKFEGERAFKAWAVARLSORPPKAEFAEVSKLVTDLTAKV 241  
Db 207 LDALKEKALVSAVRQLKCSSMKKFGERAFKAWAVARMSQTFPNADFAEITKLATDLTKV 266  
QY 242 HTECHGDLLECCADRDALAKYICENODSISSKLKECCPELLEKSHCIAEVENDEMPAD 301  
Db 267 TOECHGDLLECCADRDALAKYICENODSISSKLKECCPELLEKSHCIAEVEHDDMPAD 326  
QY 302 LPSLAADFEVSKDCKNYAEAKDVFGLMFLEYEARHPDYSVVLRLRLAKTYETTLKCC 361  
Db 327 LPALTADFEVDCKNYAEAKDVFGLTFLEYEARHPDYSVVLRLRLAKTYETTLKCC 386  
QY 362 AAADPHCYAKVDFEFPKLVPEPQNLIKQNCLEFQELGEYFQNALVRYTKKQVSTP 421  
Db 387 AEADPHACYGHVDFEFPKLVPEPQNLVKSNCLEFQELGEYFQNALVRYTKKQVSTP 446  
QY 422 TLVEVSRNLGKVGKCKKHPKAPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCTESL 481  
Db 447 TLVEAARSLGRVGHCCALPEKRUFCVEDYLSALNRLVCLLHEKTPVSEQVTKCCSGL 506  
QY 482 VNRPCFSALEVDYETVVPKFNATFTTFHADICTLSEKERQIKKOTALVLEKHKPKATK 541  
Db 507 VERRECFSALEVDYETVVPKFNATFTTFHADICTLSEKERQIKKOTALVLEKHKPKATK 566  
QY 542 EQLKAVMDDFAAFEKCKKADKDKTCFAEBGKLVAAASQAAL 583  
Db 567 EQLKAVMDDFAAFEKCKKADKDKTCFAEBGKLVAAASQAAL 608

## RESULT 17

QEB320  
ID Q6B320 PRELIMINARY; PRT; 583 AA.  
AC Q6B320;  
DT 25-OCT-2004 (Tremblrel. 28, Created)  
DT 25-OCT-2004 (Tremblrel. 28, Last sequence update)  
DT 25-OCT-2004 (Tremblrel. 28, Last annotation update)  
DE Serum albumin (Fragment).  
GN Name=ESA;  
OS Elephas maximus (Indian elephant).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Proboscidea; Elephantidae; Elephas.  
OX NCBI\_TaxID=9783;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Lazar J., Rasmussen B., Greenwood D.R., Bang I.-S., Prestwich G.D.;  
RT "Elephant Albumin: A Multi-purpose Pheromone Shuttle.";  
RL Chem. Biol. 0:0-0(2004).  
DR EMBL; AY684122; AAT90502.1; -;  
DR GO; GO:0005615; C:extracellular space; IEA.





[6]  
SEQUENCE OF 25-44.  
RC TISSUE=Liver;  
RX MEDLINE=93162044; PubMed=1286668;  
RA Giometti C.S., Taylor J., Tollaksen S.L.;  
RT "Mouse liver protein database: a catalog of proteins detected by two-dimensional gel electrophoresis."  
RL Electrophoresis 13:970-991(1992).  
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good binding capacity for water, Ca(2+), Na(+), K(+), fatty acids, hormones, bilirubin and drugs. Its main function is the regulation of the colloidal osmotic pressure of blood.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Plasma.  
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.  
CC -!- SIMILARITY: Contains 3 albumin domains.  
CC -----  
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AJ011413; CA009617.1; -;  
DR EMBL; AK010025; BAB26650.1; -;  
DR EMBL; BC024643; AAH24643.1; -;  
DR EMBL; BC049971; AAH49971.1; -;  
DR EMBL; M16111; AAA37190.1; -;  
DR EMBL; X13060; CA031458.1; -;  
DR PIR; A05139; A05139.  
DR HSP; P02768; IHKI.  
DR SWISS-2DPAGE; P07724; MOUSE.  
DR MGD; MGI:87991; Alb1.  
DR InterPro; IPR001703; Alphafoetoprot.  
DR InterPro; IPR000264; Serum albumin.  
DR Pfam; PF00273; Serum albumin; 3.  
DR PRINTS; PR00802; SERUMALBUMIN.  
DR ProDom; PD002486; Serum albumin; 1.  
DR PROSITE; PS00212; ALBUMIN; 3.  
DR Copper; Direct protein sequencing; Lipid-binding; Metal-binding;  
KW Repeat; Signal.  
FT SIGNAL 1 18 By similarity.  
FT PROPEP 19 24  
FT CHAIN 25 608 Serum albumin.  
FT DOMAIN 25 205 Albumin 1.  
FT DOMAIN 212 397 Albumin 2.  
FT DOMAIN 404 595 Albumin 3.  
FT METAL 27 27 Copper.  
FT DISULFID 77 86 By similarity.  
FT DISULFID 99 115 By similarity.  
FT DISULFID 114 125 By similarity.  
FT DISULFID 148 193 By similarity.  
FT DISULFID 192 201 By similarity.  
FT DISULFID 224 270 By similarity.  
FT DISULFID 269 277 By similarity.  
FT DISULFID 289 303 By similarity.  
FT DISULFID 302 313 By similarity.  
FT DISULFID 340 385 By similarity.  
FT DISULFID 384 393 By similarity.  
FT DISULFID 416 462 By similarity.  
FT DISULFID 461 472 By similarity.  
FT DISULFID 485 501 By similarity.  
FT DISULFID 500 511 By similarity.  
FT DISULFID 538 583 By similarity.  
FT DISULFID 582 591 By similarity.  
FT CONFLICT 27 27 H -> D (in Ref. 6).  
FT CONFLICT 33 33 H -> D (in Ref. 6).  
FT CONFLICT 41 41 Q -> I (in Ref. 6).  
SQ SEQUENCE 608 AA; 68692 MW; 2927FC7EED3A61B4 CRC64;  
Query Match 62.8%; Score 2378; DB 1; Length 608;

Best Local Similarity 72.4%; Pred. No. 8.5e-52;  
Matches 422; Conservative 80; Mismatches 81; Indels 0; Gaps 0;  
QY 1 DAHSEVAHREFKOLGEENFKALVLIAPAYLQOCPPFEDHVKLVNVEVTEFAKTCVADESAAE 60  
DB 25 EAHKSEIAHRYNDLGEQHFGLVLIAPESQYLQKCSYDEHAKLVQEVTFDFAKTCVADESAA 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETGEMADCAKQEPERNECFLOHKDNDPMLPRVVRREV 120  
DB 85 NCDKSLHTLFGDKLCAIPNLRENYGELADCCTKQEPERNECFLOHKDNDPMLPRVVRREV 144  
QY 121 DVMTAFHDNEETFLKYLVEIARRHPVYFAPELLYIAEQLYIAEQINIEITQCAEADKESCLTP 180  
DB 145 EAMCTSFKENPTTFMGLYHVARRRHPYFAPELLYIAEQLYIAEQINIEITQCAEADKESCLTP 204  
QY 181 KLDELRLDEGKASSAKQRLKCLASLQKFGERAFKAWAVARLSQRPFAEVSRLVDTLTK 240  
DB 205 KLDGKVKALVSSVVRQMKCSMQKFGERAFKAWAVARLSQTFPNADFAEITKATLTK 264  
QY 241 VHTCCHGDLLECADDRAADLAKYICENQDSISSKJKECEKPLLEKSHCIAEVDENWPA 300  
DB 265 VNKECCHGDLLECADDRAELAKYMCENQATISSKJLQTCDCPKLLKKAHCLSEVEHDTWPA 324  
QY 301 DLPSLAADFVESKOVCKNYAEAKOVFLGMLFLEYARRHPDYSVLTLRLAKTYETTLK 360  
DB 325 DLPAIADFVEDQEVCKNYAEAKOVFLGMLFLEYARRHPDYSVLTLRLAKTYETTLK 384  
QY 361 CAADAPHECYAKVDFDEKPLVEEPONLIKQNCLEFQELGEYKFKQNALVRYTKVPQVST 420  
DB 385 CAEANPPACYGTVLAEFQPLVEEPONLVKNCVLEKGLGEYGFQNALVRYTKVPQVST 444  
QY 421 PTLVEVSRLNGKVGSKCKKHPKAMPKCAEDYLSVLNQLCVLHKEKTPVSDRVTKCTES 480  
DB 445 PTLVEAARNLGRVGTCKCTLPEDQRLPCVEDYLSAILNRCVLLHKEKTPVSEHVTCKSGS 504  
QY 481 LVNRPFCPSALEVDETYYPKEFNATFTFHADICTLSEKEROIKKQTALVELVKKPKAT 540  
DB 505 LVRRPFCPSALTVDETYYPKEFNATFTFHSDICTLPEKEKQIKKQTALVELVKKPKAT 564  
QY 541 KEOLKAVMDDFAAVFEKCKKADDDKTCFAEKGKLVAAASQAAL 583  
DB 565 AEQLKTVMDDFAAFLDTCCKADKCTCFSTGPNLVTRCKDAL 607  
RESULT 19  
Q8C7H3 PRELIMINARY; PRT; 608 AA.  
ID Q8C7H3  
AC Q8C7H3; 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Mus musculus adult male liver tumor cDNA, RIKEN full-length enriched library, clone:C73003P03 product:albumin 1, full insert  
DE sequence.  
GN Name=Alb1;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Liver;  
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
RA Carninci P., Hayashizaki Y.  
RT "High-efficiency full-length cDNA cloning."  
RL Meth. Enzymol. 303:19-44(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Liver;  
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
RA RIKEN FANTOM Consortium;  
RT "Functional annotation of a full-length mouse cDNA collection."  
RL Nature 409:685-690(2001).



```
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama S., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki K.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsunura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura K.,
RA Nishi K., Nomura K., Nunazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh K., Sakai K., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK050644; BAC34360.1; -.
DR HSSP; P02768; IHKL.
DR MGD; MGI:87991; Alb1.
DR GO; GO:0005615; C:extracellular space; TAS.
DR InterPro; IPR001703; Alphafetoprot.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; Serum albumin; 3.
DR PRINTS; PR00803; AFETOPROTEIN.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
FT NON_TER
FT SEQUENCE 576 AA; 65002 MW; F85733899AE37F04 CRC64;
Query Match 61.7%; Score 2336; DB 2; Length 576;
Best Local Similarity 72.2%; Pred. No. 8.4e-51;
Matches 415; Conservative 79; Mismatches 81; Indels 0; Gaps 0;
QY 9 HRFKDLGENFKALVLIAPQYLQCFEDHVKLVNEVTEPAKTCVADSAENCKSLHT 68
DB 1 NRYNDLGEHFHGLVLIAPQYLQCFEDHVKLVNEVTEPAKTCVADSAENCKSLHT 60
QY 69 LFQDKLCTVATRETYGENADCCAKOEPERNECFLOHKDDNPNLRLVPEVDVMTAFH 128
DB 61 LFQDKLCAIPNRENYGENADCCAKOEPERNECFLOHKDDNPNLRLVPEVDVMTAFH 120
QY 129 DNEETFKKLYEIAARRHPVYAPPELLFFAKRYKAAFTBCCQAADKAAACLLPKLDELRLDE 188
DB 121 ENPTTFMGVYLHEVARRHPVYAPPELLFYAEQYNELTQCCAEADKESCLTPKLDGVKEK 180
QY 189 GKASSAKQRLKASLQKQGERAPKAWARLSQRFPAKSAFVSKVLDTLTKVHTCECHG 248
DB 181 ALVSVVRQRMKSSMQKGERAPKAWARLSQTFPNADFAETIKLATDLTKVNEKCECHG 240
QY 249 DLLECADDRADIATKYICENQDSTSSKLKSCCKEPLKSHCIAEVENDEMPADLPISLAAD 308
DB 241 DLLECADDRAEALAKYCNENQATISSKLTQCCDKPKLLKKAHCUSEVEHDTMPADLPISLAAD 300
```

```
QY 309 FVESKDVCKYNAEAKDVPGLMFLYEHARRHPDYSVILLRLAKTYETTTLEKCCAAADPHE 368
DB 301 FVEDPECKYNAEAKDVPGLMFLYEHARRHPDYSVILLRLAKTYETTTLEKCCAAADPHE 360
QY 369 CYAKVDFDKPLVEEPQNLKQNCFLPQLEGEYKFNQALLVRYTKVPQVSTPTTIVESVR 428
DB 361 CYGVTLAEFQPLVBEPKNLVNTCDLYEKLGEYGFQNALVRYTKQKAPQVSTPTTIVEAAR 420
QY 429 NLGVGSKCCXHPHAKRMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCTCTSLVNRPCF 488
DB 421 NLGRVGTCKCTLPEDQRLPCVEDYLSALNLRVCLLHEKTPVSEHVTKCCSGSLVRRPCF 480
QY 489 SALEVDETVYKPEFNAETFTSHADICTLSEKEROIKKOTALVELVHKHPKATKEOLKAVM 548
DB 481 SALTVDVETVPKPEKFAETFTTHSDICTLPSKEKQIKKOTALAEVLVHKHPKATKEOLKAVM 540
QY 549 DDFAAFEKCKCKADDKETCFAEEGKKLVAAASQAAL 583
DB 541 DDFAQFLDTCCCKADKDTCFSTEGPNLVTRCKDAL 575
QY 549 DDFAAFEKCKCKADDKETCFAEEGKKLVAAASQAAL 583
DB 541 DDFAQFLDTCCCKADKDTCFSTEGPNLVTRCKDAL 575
RESULT 21
Q86YGO PRELIMINARY; PRT; 417 AA.
AC Q86YGO;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ALB protein.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC041789; AAH41789.1; -.
DR HSSP; P02768; 1E7B.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR00264; Serum albumin.
DR Pfam; PF00273; Serum albumin; 2.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 2.
DR PROSITE; PS00212; ALBUMIN; 2.
```



ALBU CHICK  
 ID ALBU CHICK STANDARD; PRT; 615 AA.  
 AC P19121;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 05-JUN-2004 (Rel. 44, Last annotation update)  
 DE Serum albumin precursor (Alpha-livetin) (Allergen Gal d 5).  
 GN Name=ALB;  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OC NCBI\_TaxID=9031;  
 RN (1)  
 RN SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RA Cassidy A.I., Salklilid C.K., Baverstock P., Wallace J.C.;  
 RL Submitted (JUL-1991) to the EMBL/GenBank/DBJ databases.  
 RN (2)  
 RN SEQUENCE OF 1-28 FROM N.A.  
 RX MEDLINE=83161037; PubMed=6187737;  
 RA Hache R.J.G., Wiskocil R., Vasa M., Roy R.N., Lau P.C.K., Deeley R.G.;  
 RT "The 5' noncoding and flanking regions of the avian very low density  
 RT apolipoprotein II and serum albumin genes. Homologies with the egg  
 RT white protein genes.";  
 RL J. Biol. Chem. 258:4556-4564(1983).  
 RN (3)  
 RN SEQUENCE OF 19-30.  
 RX MEDLINE=78019943; PubMed=911327;  
 RA Rosen A.M., Geller D.M.;  
 RT "Chicken microsomal albumin: amino terminal sequence of chicken  
 RT proalbumin.";  
 RL Biochem. Biophys. Res. Commun. 78:1060-1066(1977).  
 RN (4)  
 RN ALLERGENIC PROPERTIES.  
 RX MEDLINE=21381307; PubMed=11488669;  
 RA Quirce S., Maranon F., Umplierrez A., de las Heras M.,  
 RA Fernandez-Caldas E., Sastre J.;  
 RT "Chicken serum albumin (Gal d 5\*) is a partially heat-labile inhalant  
 RT and food allergen implicated in the bird-egg syndrome.";  
 RL Allergy 56:754-762(2001).  
 CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good  
 CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
 CC hormones, bilirubin and drugs. Its main function is the regulation  
 CC of the colloidal osmotic pressure of blood.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Plasma.  
 CC -!- ALLERGEN: Causes an allergic reaction in human. Binds IgE.  
 CC Partially heat-labile allergen that may cause both respiratory and  
 CC food-allergy symptoms in patients with the bird-egg syndrome.  
 CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.  
 CC -!- SIMILARITY: Contains 3 albumin domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X60688; CAA43098.1; -;  
 DR EMBL; V00381; CAA23680.1; -;  
 DR PIR; S15571; ABCS.  
 DR HSSP; P02768; 1HK1.  
 DR InterPro; IPR001703; Alphafetoprot.  
 DR InterPro; IPR000264; Serum albumin.  
 DR Pfam; PF00273; Serum albumin; 3.  
 DR PRINTS; PR00802; SERUMALBUMIN.  
 DR ProDom; PD002486; Serum albumin; 1.  
 DR SMART; SM00103; ALBUMIN; 3.  
 DR PROSITE; PS00212; ALBUMIN; 3.  
 KW Allergen; Copper; Direct protein sequencing; Lipid-binding;

Metal-binding; Repeat; Signal.  
 KW SIGNAL 1 18  
 FT PROPEP 19 23  
 FT CHAIN 24 615  
 FT DOMAIN 24 209  
 FT DOMAIN 216 401  
 FT DOMAIN 408 599  
 FT METAL 30 30  
 FT DISULFID 80 89  
 FT DISULFID 102 118  
 FT DISULFID 117 128  
 FT DISULFID 152 197  
 FT DISULFID 196 205  
 FT DISULFID 228 274  
 FT DISULFID 273 281  
 FT DISULFID 293 307  
 FT DISULFID 306 317  
 FT DISULFID 344 389  
 FT DISULFID 388 397  
 FT DISULFID 420 466  
 FT DISULFID 465 476  
 FT DISULFID 489 505  
 FT DISULFID 504 515  
 FT DISULFID 542 587  
 FT DISULFID 586 595  
 FT CARBOHYD 500 500  
 FT CONFLICT 24 24  
 SQ SEQUENCE 615 AA; 69918 MW; 559E4BBAC066C6 CRC64;  
 Query Match 41.2%; Score 1557.9; DB 1; Length 615;  
 Best Local Similarity 46.7%; Pred. No. 6.7e-31;  
 Matches 273; Conservative 118; Mismatches 19; Indels 1; Gaps 1;  
 QY 3 HKSEVAHFRKDLGGENFKALVLIAPQYLOQCPEFDHVKLVNEVTEFAKTCVADSAENC 62  
 DB 30 HKSIAHRYNDLKEETFKAVAMITFAQYLCRSYEGSLKLVKDVVDLQAKCVANEDAPC 89  
 QY 63 DKSHTLFGDKLCTVATLRETYGEMADCCAKQEBERNECFLOHKDDNPNLPR-LVRPEVD 121  
 DB 90 SKPLPSIILDRICQVEKLKRDYSGAMADCCSKADPERNECFLSFKVSPDFVQPYQRPASD 149  
 QY 122 VMCTAFHNEETELKYLIEIARRHPYFVAPPELLAFKRYKAAFTCCQAAADKAAACLIPK 181  
 DB 150 VICQEQDNRFSLGHFIYSVARHPFLYPAILSFVDFEHALQSCCKESDVGACLDTK 209  
 QY 182 LDELDECKASAKQRLKASLQKFGERAFAKAWAVARLSQRFPFAFVSKLVLTDLTKV 241  
 DB 210 EIVMREKAGVSVKQYFCGILKQFGRVFAQRLIYLSQYKAPFSEVSKFVHDSIGV 269  
 QY 242 HTECCHGOLLLECADDRADLAKYICENQDISISKLECKECPLEKSHCIAEVENDEMPAD 301  
 DB 270 HKECCGDMVECDMDARMMSNLCSQDVFSGIKIDCKCEKPIVERSQCIIMEAFDEKPAD 329  
 QY 302 LPSLAADFVESKDVCKENYAEAKDVFGLMFVEYARRHPDYSVVLLRLAKTVETTLKCC 361  
 DB 330 LPSLVKEVIEDKECKSEAGHDAPFAEFVYSRRHPEFSIQLIMRLAKGVESLLEKCC 389  
 QY 362 AAADPHECYAKVDFDKPLVVEPQNLIKQNCLEFQLGEGYKFNALLVRYTKKVPQVSTP 421  
 DB 390 KTDNPAECYANAOEQNLHIKETQDVKTNCDDLHDHGEADFLKSLIRYTKKMPQVPTD 449  
 QY 422 TLVEVSRNLGKVGSKCKHPKAPKMPCAEDYLVSVVNLQICVLHEKTPVSDRVTKCTTESL 481  
 DB 450 LLELTGKKMTTIGTKCCQLGEDRRMACSEGYLSIVIHDTCRKQETTPINDNVSQCCSOY 509  
 QY 482 VNRRCFSALEVDETYVPEKFAETFTFHADICTLSEKERQIKQTALVELVKHKPKATK 541  
 DB 510 ANRRECFAMGVDTKYVPPPPNPDMFSEFDEKCSAPABEREVGQMKLINLINLRKRPQMT 569  
 QY 542 EQLKAVMDDFAAFVEKCKCKADDDKTCFAEEGKKLVAAASQAALGL 585  
 DB 570 EQIKTIADGFTAMVDKCKQSQDINTCFEGEGANLIVQSEATLGI 613



## RESULT 24

```
Q8JIA9 PRELIMINARY; PRT; 527 AA.
AC Q8JIA9;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Serum albumin (Fragment).
OS Sphenodon punctatus (Chordata) (Tuatara).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Sphenodontia; Sphenodontidae; Sphenodon.
OX NCBI_TaxID=8508;
RN [1]
SEQUENCE FROM N.A.
RA Metcalf V.J., Brennan S.O., George P.M., Chambers G.K.;
RP Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF375971; AA046104.1; -.
DR HSP; P02768; I27B.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR001703; Alpha-fetoprot.
DR InterPro; IPR00264; Serum albumin.
DR Pfam; PF00273; Serum albumin; 3.
DR PRINTS; PR00803; Serum albumin; 3.
DR PRODOM; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
FT NON TER 1
SQ SEQUENCE 527 AA; 59711 MW; C62B7998387F5929 CRC64;

Query Match 34.3%; Score 1296.7; DB 2; Length 527;
Best Local Similarity 45.9%; Pred. No. 1.9e-24;
Matches 240; Conservative 92; Mismatches 188; Indels 3; Gaps 3;

Qy 62 CDKSLHTLFDGKLTCTVATLRETYGEMADCAKQPERNECFQHKDQNP-N-LRPLVRPEV 120
Db 5 CLKSLDTIFLDEICHEEGFAKY-DLAACCAKAEVERKECLLAHKNAATPGFIFAFQPGI 63

Qy 121 DVMTAFDNEETPLKVLVEIARRHPYFVAPELLFFAKYKAFTCCQAADKAACLLP 180
Db 64 EVSKULQDDRLLTLGNYIYEVARHPYLQVPVFATSLYDEALTKCCQTADKATCFHP 123

Qy 181 KLDELDEGKASSAKQRLKASLQKGERAFKAWARLSQRPFKABFAEVSKLVTDLTK 240
Db 124 RIPPLIEVLKWSNGIQENTCGILKKGERTLKATKLQMSQKPKADFATINKLVEDITH 183

Qy 241 VHTCCGGDLLECAADRADLAKYICENODSISSKLECCBKPLLEKSHCHIAEVNDMPA 300
Db 184 MHTCCRGDTLECLRDREALTEYTCSHKDAISSKLPCTCCESVLERGECIVRLNDDKPA 243

Qy 301 DLPSLAADFVESKDVCNVAEAKDVFGLMFLYEYARHPDYSVVLRLAKTVETTLK 360
Db 244 DLSERIAEYEDHPVCDHLAKEQDAFLAKFLYEYSRRHPPELSTQILGVGKGQYQLLERC 303

Qy 361 CAADAPHECAKVPDEFKPLVEEPQNLIKONCELFEOQLGEYKFQNALLVRYTKVPQVST 420
Db 304 KKTNDPPECYQGAEDLKKHIAQFQELVQNCDLNTLGLYLFHALLIYRTKEMPQLTS 363

Qy 421 PTLVEVRNGLKGVSKCKGKPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCTES 480
Db 364 EELIIFYTR-ITKAASRCCEVSVDKLPCTEGYDFVLGQICQHRQSSINNVNQCQCSN 422

Qy 481 LVNRRPCFSALEVDETVVPKEFNATETFEADICTLSEKEROIKQATALVELVKHKPKAT 540
Db 423 YALRSLCITSLGGDEKVPLEFSDLFTTFHEDLCHAAQDKLQERKQMI VNLVKHKNIT 482

Qy 541 KEQLKAYMDDFAAEVKECKADKCTCFABEGKKLVAASQAAL 583
Db 483 KEQLQTVFGGFTKWKTEKCCRAEDHEACFGEGGPKLVAESQAL 525
```

## RESULT 25

```
FETA_PANTR STANDARD; PRT; 609 AA.
AC Q28789;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 03-JUL-2004 (Rel. 44, Last annotation update)
DE Alpha-fetoprotein precursor (Alpha-fetoglobulin) (Alpha-1-fetoprotein).
DE Name=AFP;
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=96032345; PubMed=7557431; DOI=10.1016/0378-1119(95)00303-N;
RA Nishio H., Gibbs P.E., Minghetti P.P., Zielinski R., Dugaiczky A.;
RT "the chimpanzee alpha-fetoprotein-encoding gene shows structural similarity to that of gorilla but distinct differences from that of human.";
RL Gene 162:213-220(1995).
CC -I- FUNCTION: Binds copper, nickel, and fatty acids as well as, and bilirubin less well than, serum albumin.
CC -I- SUBUNIT: Dimeric and trimeric forms have been found in addition to the monomeric form (By similarity).
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- TISSUE SPECIFICITY: Plasma. Synthesized by the fetal liver and yolk sac.
CC -I- SIMILARITY: Belongs to the ALB/AFP/VDB family.
CC -I- SIMILARITY: Contains 3 albumin domains.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
-----
EMBL; U21916; AAA91641.1; -.
DR PIR; JC4258; JC4258.
DR HSP; P02768; I27B.
DR InterPro; IPR001703; Alpha-fetoprot.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; Serum albumin; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR PRODOM; PD002486; Serum albumin; 1.
DR PROSITE; PS00212; ALBUMIN; 2.
KW Copper; Glycoprotein; Metal-binding; Nickel; Repeat; Signal; Sulfation.
FT SIGNAL 1 18 By similarity.
FT CHAIN 19 609 Alpha-fetoprotein.
FT DOMAIN 20 205 Albumin 1.
FT DOMAIN 212 397 Albumin 2.
FT DOMAIN 404 595 Albumin 3.
FT METAL 22 22 Copper and nickel (By similarity).
FT DISULFID 99 114 By similarity.
FT DISULFID 113 124 By similarity.
FT DISULFID 148 193 By similarity.
FT DISULFID 192 201 By similarity.
FT DISULFID 224 270 By similarity.
FT DISULFID 269 277 By similarity.
FT DISULFID 289 303 By similarity.
FT DISULFID 302 313 By similarity.
FT DISULFID 384 393 By similarity.
FT DISULFID 416 462 By similarity.
FT DISULFID 461 472 By similarity.
FT DISULFID 485 501 By similarity.
FT DISULFID 500 511 By similarity.
FT DISULFID 538 583 By similarity.
FT DISULFID 582 591 By similarity.
```



FT CARBOHYD 42 42 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 251 251 N-linked (GlcNAc...) (Potential).  
SQ SEQUENCE 609 AA; 68741 MW; C032987CAD0E672B CRC64;

Query Match 33.2%; Score 1256.7; DB 1; Length 609;  
Best Local Similarity 39.9%; Pred. No. 3.1e-23;  
Matches 237; Conservative 113; Mismatches 231; Indels 13; Gaps 3;

Qy 4 KSEVAHR -----FKDLGSENFALVLIAPAOYLQCPEDHVKLVNEVTEPAKT 52  
Db 17 ESRTLHRENYGIASLDSYQCTAEINLTDLATIPFAQVQSEATYKESKMKVDALTAIEK 76  
Qy 53 CVADESANCKSLHTLFGDKLCTVATRETYGEMADCAQBEPRNECFLOHODNP-N 111  
Db 77 PTGDEQSAGCLENQLPAFLLELCREKTEIKYGH-SDCCSQSEBGRHNCFLAHKPTPAS 135  
Qy 112 LPLRVPRVDVMTAFHNEFTLKKLYETARRHPYVAPPELLAFKRYKAAFTCCQA 171  
Db 136 IPFQVPEVTSCEAYEDRETFFMKFYIETARRHPFLYAPITLLWAARYDKIIPSCCKA 195  
Qy 172 ADKAACLLPKLDELDEGKASSAKORLKCASLQKGERAFKAWAVARLSORFPKAFPAEV 231  
Db 196 ENAVECFQTKAATVTKELRESSLLNQHACVMKNFGTRTFOAITVTKLSQKFTKVNFEI 255  
Qy 232 SKLVTDLTQVTECHGDLLECADRADLAKYICENQDSISKLKCECEKPLEKSHCIA 291  
Db 256 OKLVLDVAHVHEHCCRGDVLDCLODGEKIMSYCSQDPTLSNKITECCKLTTLERGQCII 315  
Qy 292 EVENDEMDADLPSLAADFVESKDVCKNYAEAKDVFGLMFLYERHDPDYSVLLLRKAK 351  
Db 316 HAENDEKPEGLSPNLRFLGDRDFNQFSGSEKNIFLAGFVHYSHRHPQLAVSVILRVAK 375  
Qy 352 TVETLEKCAAAAPHECVAKVDFDKPLVBEPPQMLIKONCELPQLGEYFQNALVRY 411  
Db 376 GYQELLEKCFQTEPLEQDQGEELQYIQESQALAKRSGLFKLGEYIYQNAFLVAY 435  
Qy 412 TKKVPQVSTPTLVEVSRNLKGVSKCKHPEAKRMPCAEDYLSVVLNOLCVLHEKTPVSD 471  
Db 436 TKKAPQTSSELMAITRMAATAATCCQLSEDKLLACGEGAADIIGHLCIRHETTPVNP 495  
Qy 472 RYTKCTESLVNRRCFSALEVDYVPKFEAFETFTHADICTLSEKERQIKKQTALVE 531  
Db 496 GVGQCCTSSYANRRFCFSLVVDYVPPAFSDDKPEIFHKDLCQAGQVALQTMKQEFLLN 555  
Qy 532 LVKHKPKATKEOLKAVMDFAFVEKCKKADDKETCFABEGKLVAAASQALGL 585  
Db 556 LVKQRPQITEQEAVIAVDFSGLEKCCQGEVCFABEGQKLISKTFEALGV 609

RESULT 26  
FETA HUMAN  
ID FETA HUMAN STANDARD; PRT; 609 AA.  
AC P02771;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Alpha-fetoprotein precursor (Alpha-fetoglobulin) (Alpha-1-fetoprotein).  
DE Name=AFP;  
GN Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=83273664; PubMed=6192439;  
RA Morinaga T., Sakai M., Wegmann T.G., Tamaoki T.;  
RT "Primary structures of human alpha-fetoprotein and its mRNA";  
RL Proc. Natl. Acad. Sci. U.S.A. 80:4604-4608 (1983).  
RN [2]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=87185438; PubMed=2436661;  
RA Gibbs P.E.M., Zielinski R., Boyd C., Dugaiczky A.;

"Structure, polymorphism, and novel repeated DNA elements revealed by a complete sequence of the human alpha-fetoprotein gene."; Biochemistry 26:1332-1343 (1987).  
[3]  
RN SEQUENCE FROM N.A.  
RP TISSUE=Lung;  
RC MEDLINE=2238257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F., Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullihy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."; Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
[4]  
RN SEQUENCE OF 1-28 FROM N.A.  
RP MEDLINE=93278385; PubMed=7684942;  
RA McVey J.H., Michaelides K., Hansen L.P., Ferguson-Smith M., Tilghman S., Krumlauf R., Tuddenham E.G.D.;  
RT "A G->A substitution in an HNF I binding site in the human alpha-fetoprotein gene is associated with hereditary persistence of alpha-fetoprotein (HPAFP)"; Hum. Mol. Genet. 2:379-379 (1993).  
[5]  
RN SEQUENCE OF 429-556 FROM N.A.  
RP MEDLINE=83158778; PubMed=6187626; DOI=10.1016/0378-1119(82)90210-4;  
RA Beattie W.G., Dugaiczky A.;  
RT "Structure and evolution of human alpha-fetoprotein deduced from partial sequence of cloned cDNA."; Gene 20:415-422 (1982).  
[6]  
RN PARTIAL SEQUENCE OF 19-609.  
RP MEDLINE=91242409; PubMed=1709810;  
RA Pucci P., Siciliano R., Malorni A., Marino G., Tecce M.F., Ceccarini C., Terrana B.;  
RT "Human alpha-fetoprotein primary structure: a mass spectrometric study."; Biochemistry 30:5061-5066 (1991).  
[7]  
RN PRELIMINARY SEQUENCE OF 19-35.  
RP MEDLINE=77242506; PubMed=70228; DOI=10.1016/0005-2795(77)90198-2;  
RA Yachnin S., Hsu R., Heinrichson R.L., Miller J.B.;  
RT "Studies on human alpha-fetoprotein. Isolation and characterization of monomeric and polymeric forms and amino-terminal sequence analysis."; Biochim. Biophys. Acta 493:418-428 (1977).  
[8]  
RN PRELIMINARY SEQUENCE OF 19-38.  
RP MEDLINE=78001760; PubMed=711198;  
RA Aoyagi Y., Ikenaka T., Ichida F.;  
RT "Comparative chemical structures of human alpha-fetoproteins from fetal serum and from ascites fluid of a patient with hepatoma."; Cancer Res. 37:3663-3667 (1977).  
[9]  
RN PRELIMINARY SEQUENCE OF 19-39.  
RP MEDLINE=75018719; PubMed=4138095;  
RA Ruoslahti E., Pihko H., Vaheri A., Seppala M., Virolainen M., Kontinen A.;  
RT "Alpha fetoprotein: structure and expression in man and inbred mouse strains under normal conditions and liver injury."; Johns Hopkins Med. J. Suppl. 3:249-255 (1974).

RA [10] GENE STRUCTURE.  
RP MEDLINE=85182629; PubMed=2580830;  
RX Sakai M., Morinaga T., Urano Y., Watanabe K., Wegmann T.G.,  
RA Tamaoki T.;  
RT "The human alpha-fetoprotein gene. Sequence organization and the 5'  
RT flanking region.";  
RL J. Biol. Chem. 260:5055-5060(1985).  
RN [11]  
RP METAL-BINDING.  
RX MEDLINE=79001617; PubMed=80265;  
RA Aoyagi Y., Ikenaka T., Ichida F.;  
RT "Copper(II)-binding ability of human alpha-fetoprotein.";  
RL Cancer Res. 38:3483-3486(1978).  
RN [12]  
RP BILIRUBIN-BINDING.  
RX MEDLINE=80001710; PubMed=89900;  
RA Aoyagi Y., Ikenaka T., Ichida F.;  
RT "Alpha-Fetoprotein as a carrier protein in plasma and its bilirubin-  
RT binding ability.";  
RL Cancer Res. 39:3571-3574(1979).  
RN [13]  
RP SULFATION.  
RX MEDLINE=86042625; PubMed=2414772;  
RA Liu M.C., Yu S., Sy J., Redman C.M., Lipmann F.;  
RT "Tyrosine sulfation of proteins from the human hepatoma cell line  
RT HepG2.";  
RL Proc. Natl. Acad. Sci. U.S.A. 82:7160-7164(1985).  
CC -!- FUNCTION: Binds copper, nickel, and fatty acids as well as, and  
CC bilirubin less well than, serum albumin. Only a small percentage  
CC (less than 2%) of the human AFP shows estrogen-binding properties.  
CC -!- SUBUNIT: Dimeric and trimeric forms have been found in addition to  
CC the monomeric form.  
CC -!- TISSUE SPECIFICITY: Plasma. Synthesized by the fetal liver and  
CC yolk sac.  
CC -!- DEVELOPMENTAL STAGE: Occurs in the plasma of fetuses more than 4  
CC weeks old, reaches the highest levels during the 12th-16th week of  
CC gestation, and drops to trace amounts after birth. The serum level  
CC in adults is usually less than 40 ng/ml. AFP occurs also at high  
CC levels in the plasma and ascitic fluid of adults with hepatoma.  
CC -!- PTM: Independent studies suggest heterogeneity of the amino-  
CC terminal sequence of the mature protein and of the cleavage site  
CC of the signal sequence.  
CC -!- PTM: Sulfated.  
CC -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.  
CC -!- SIMILARITY: Contains 3 albumin domains.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M10949; AA51674.1; -;  
CC EMBL; M10950; AA51675.1; -;  
CC EMBL; V01514; CAA24758.1; -;  
CC EMBL; M16110; AAB58754.1; -;  
CC EMBL; BC027881; AAB57881.1; -;  
CC EMBL; Z19532; CAA79592.1; -;  
CC PIR; A26624; FPHU.  
CC HSSP; P02768; 1E7B.  
CC GlycoSuiteDB; P02771; -;  
CC Siena-2DPAGE; P02771; -;  
CC Genew; HGNC:317; AFP.  
CC H-InVDB; HIX0004279; -;  
CC MIM; 104150; -;  
CC InterPro; IPR001703; Alphafetoprot.  
CC InterPro; IPR000264; Serum albumin.  
CC Pfam; PF00273; Serum albumin; 3.  
CC PRINTS; PR00802; SERUMALBUMIN.

DR ProDom; PD002486; Serum\_albumin; 1.  
DR SMART; SM00103; ALBUMIN; 3.  
DR PROSITE; PS00212; ALBUMIN; 2.  
KW Copper; Direct protein sequencing; Glycoprotein; Metal-binding;  
KW Nickel; Polymorphism; Repeat; Signal; Sulfation.  
FT SIGNAL 1 18  
FT CHAIN 19 609 Alpha-fetoprotein.  
FT DOMAIN 20 205 Albumin 1.  
FT DOMAIN 212 397 Albumin 2.  
FT DOMAIN 404 595 Albumin 3.  
FT METAL 22 22 Copper and nickel.  
FT DISULFID 99 114  
FT DISULFID 113 124  
FT DISULFID 148 193  
FT DISULFID 192 201  
FT DISULFID 224 270  
FT DISULFID 269 277  
FT DISULFID 289 303  
FT DISULFID 302 313  
FT DISULFID 384 393  
FT DISULFID 416 462  
FT DISULFID 461 472  
FT DISULFID 485 501  
FT DISULFID 500 511  
FT DISULFID 538 583  
FT DISULFID 582 591  
FT CARBOHYD 251 251  
FT VARIANT 570 570  
FT A -> G (in dBSNP:1057173).  
FT /FTID=VAR\_012049.  
SQ SEQUENCE 609 AA; 68677 MW; 4D4E45820E1C2D4F CRC64;  
Query Match 33.1%; Score 1252.7; DB 1; Length 609;  
Best Local Similarity 39.7%; Pred No. 3.9e-23;  
Matches 236; Conservative 114; Mismatches 231; Indels 13; Gaps 3;  
Qy 4 KSEVAHR-----FKDLGEENFKALVLIAPAQYLOQCPEDHVKLVNEVTEFAKT 52  
Db 17 ESRTLHNEYGIASILDYSQCTABEISLADLATIFFAQVQEAITYKEYSKMVKDALTAIEK 76  
Qy 53 CVADESANEDKSLHTLFGKLCIVATLRETYGEMADCCAKQEPERNECFLOHDDNP-N 111  
Db 77 PTGDEQSSGGLNQLPAPLEELCHEKEILEKYGH-SDCCSQSEGRHNCFLAHKKPTPAS 135  
Qy 112 LRLVRPEVDVMTAFHDNEETFLKYLEIARHPHYFAPPELLFFAKYKAAATECCOA 171  
Db 136 IPLFQVPEPTVSCAYEEDRETFWNKFIYIARHPFLYAPTILLWAARDYDKIIPSCCKA 195  
Qy 172 ADKAAICLLPKLDELDRDEGKASSAKQRLKCSAQKQGERAFKAWAVARLSQRPFKAEFAV 231  
Db 196 ENAVECFQTKRAATVTKELRESSLLNQHACAVMKNGFTTFOAITVKLSQKFTKVNFTET 255  
Qy 232 SKLVTDLTQVHTECHGDLLECADRADLAKYICENQDSISSKUKCECEKPLLEKSHCIA 291  
Db 256 QKLVLDVAHVHEHCRCGDVLDCLQDGEKIMSYICSQODTILSNKITTECKLTTLERGGCII 315  
Qy 292 EVENDENPADLPSLAADFVSKVNAEAKDVFGLGFLYEVYARRPDYVLLLRLLAK 351  
Db 316 HAENDEKPEGLSPNLRPLGDRDNFNQSSGEKNIFLASFVHEYRRRPLQAVSVILRVAK 375  
Qy 352 TYETTTLEKCAAADPHECYAKVDFDEKPLVEEPQNLIKONCELFQELGEYKFQNALVRY 411  
Db 376 GYQELLEKCFQTNPLEQCQDGBEELQYIQESQALAKRSCGLQKLGYYLQNAFLVAY 435  
Qy 412 TKQVPQVSTPTLVEVSNLKGKCKKHPKAKMPCADYLSVVLNQLCVLHEKTPVSD 471  
Db 436 TKKAPQLTSSSELMAITRKMAATAATCCQLSDEKILACGEGAADIIGHLICIRHEMTVPNP 495  
Qy 472 RVTKCCTESLVNRRPCFSALEVDVETYPKBNAAETFTFHADICTLSEKERQIKQTALVE 531  
Db 496 GVGCCTSSSVANRRPCFSSLVVDVETIYPPAFSDDKFIHFKDLCCQAQGVALQTMKEFLIN 555  
Qy 532 LVKHKPKATKEQLKAVMDDFAAAFVEKCKCKADDDKTCFAEKGKLVVAASQAALGL 585

Db

Qy 475 KCCTESLVNRRPCFSALEVDVTVYKPEFNATETTHADICTLSEKEROIKKOTALVELVK 534  
:  
Db 508 OCCSHSLSSQTFCFSALPFDVTVPPLSVASNFNDCLTSEPQQSKKVFLIRLMK 567  
Qy 535 HKPKATLBQLKAVMDMDFAAFEVECKCCKADDKETCFABEGGKLVAASAQAALGLMSPRLEV 592  
Db 568 QYPHMTDQLTKTCVVNFVPWVDQCCKADNHNECFALLEGAKLIIDACKAILA-VHPAVEV 624

RESULT 28

FETA GORGO

ID PETA\_GORGO STANDARD; PRS; 609 AA.

AC FZ8050;

AD DT 01-AUG-1992 (Rel. 23, Created)

DT 01-AUG-1992 (Rel. 23, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE DE Alpha-fetoprotein precursor (Alpha-fetoglobulin) (Alpha-1-fetoprotein).

DE Name=AFP;

GN Gorilla gorilla gorilla (Lowland gorilla).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.  
OX NCBI\_TaxId=9595;  
RN [1]

SEQUENCE FROM N.A.

RX MEDLINE=91169517; PubMed=1706310;  
RA Ryan S.C., Zielinski R., Dugaiczak A.;  
RT "Structure of the gorilla alpha-fetoprotein gene and the divergence of  
RT primates.";  
RL Genomics 9:60-72(1991).

-I- FUNCTION: Binds copper, nickel, and fatty acids as well as, and  
CC bilirubin less well than, serum albumin.  
CC -I- SUBUNIT: Dimeric and trimeric forms have been found in addition to  
CC the monomeric form (By similarity).  
CC -I- SUBCELLULAR LOCATION: Secreted.  
CC -I- TISSUE SPECIFICITY: Plasma.  
CC -I- PTM: Sulfated (By similarity).  
CC -I- SIMILARITY: Belongs to the ALB/AFP/VDB family.  
CC -I- SIMILARITY: Contains 3 albumin domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).

EMBL; M38272; AAA73520.1; -.  
PIR; A37970; FPGO.  
HSP; P02768; IHK2.  
DR InterPro; IPR001703; Alphafetoprot.  
DR InterPro; IPR000264; Serum albumin.  
DR Pfam; PF00273; Serum albumin; 3.  
DR PRINTS; PO00802; SERUMALBUMIN.  
DR ProDom; PD002486; Serum albumin; 1.  
DR SMART; SM00103; ALBUMIN; 3.  
DR PROSITE; PS00212; ALBUMIN; 2.  
KW Copper; Glycoprotein; Metal-binding; Nickel; Repeat; Signal;  
Sulfation.

	KW	SIGNAL	FT
FT CHAIN	1	18	By similarity.
FT DOMAIN	19	609	Alpha-fetoprotein.
FT DOMAIN	20	205	Albumin 1.
FT DOMAIN	212	397	Albumin 2.
FT DOMAIN	404	595	Albumin 3.
FT METAL	22	22	Copper and nickel (By similarity).
FT DISULFID	99	114	By similarity.
FT DISULFID	113	124	By similarity.
FT DISULFID	148	193	By similarity.
FT DISULFID	192	201	By similarity.
FT DISULFID	224	270	By similarity.

```
FT DISULFID 269 277 By similarity.
FT DISULFID 289 303 By similarity.
FT DISULFID 302 313 By similarity.
FT DISULFID 384 393 By similarity.
FT DISULFID 416 462 By similarity.
FT DISULFID 461 472 By similarity.
FT DISULFID 485 501 By similarity.
FT DISULFID 500 511 By similarity.
FT DISULFID 538 583 By similarity.
FT DISULFID 582 591 By similarity.
FT CARBOHYD 251 251 N-linked (GlcNAc...) (potential).
SQ SEQUENCE 609 AA; 68697 MW; EBAE548377D960B8 CRC64;

Query Match 32.9%; Score 1245.7; DB 1; Length 609;
Best Local Similarity 39.4%; Pred. No. 5.8e-23;
Matches 234; Conservative 115; Mismatches 232; Indels 13; Gaps 3;

Qy 4 KSEVAHR-----FKDLGENFKALVLIAPFAVQLQCCPEDHVKLVNVEVTEFAKT 52
Db 17 ESRTLHRENYGIASILDSYQCTAISLADLATIFFAQFVQEAITYEYKSVKVKDALTAIEK 76
Qy 53 CVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPNERNECFLOHKDDNP-N 111
Db 77 PTGDEQAGCLENQLPAFLBELCHEKEILEKYG-LSDCSSQSEGRHNCFLAHKKPTPAS 135
Qy 112 LPLRVREVDVMTAFHDNEETFLKYLITAREHPYFYAPELLFFAKRYKAAFTCCQA 171
Db 136 IPLFQVPEPVTSCYBAYBEDRETFWNKFIETAREHPFLYAPTILLWAARVDKIIPSCCKA 195
Qy 172 ADKAACLLPKLDLDEDEGKASSAKQRLKASLQKFGERAFKAWAVARLSORFPKABFAEV 231
Db 196 ENAVECQTKAATVKELRESSLLNQHACAVKNGFGTRTQAIVTKLSQKFKVNTETI 255
Qy 232 SKLVTDLTKVHTECHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIA 291
Db 256 QKLVLDVAHVEHCRCGVDLCLDQGEKINSYICSQQDTLSNKITECKLTTLERGQCII 315
Qy 292 EVENDEMPADLPSLAADFVESKDKVYAEAKVDFGLMFLYEHARRHPDYSVVLLRLAK 351
Db 316 HAENDEKPEGLSPNLNRDGRDNQFSSGKNIFLASFVHEYSRRHPQLAVSVILKVAK 375
Qy 352 TYETTLKCCAAADPHECYAKVDFEFPVLEBPQNLIKONCELFQELGEYKFNALLVRY 411
Db 376 GYQELLECKCFQTEPLECQDKGEELQYIQESQALAKRSCGLFQKLGYYLQNAFLVAY 435
Qy 412 TKKVPQVSTPTLVEVSRLNLGKVGSKCKHPKAPKMPCAEDYLSVVLNQLCVLHKEKTPVSD 471
Db 436 TKKAPQLTSSSELMAITRKMAATAATCCQLSBDKLLACGEGAADIIGHLCIRHEMTVPNP 495
Qy 472 RYTKCTTESLVNRRPCFSALEVDYVYKPEFNAETFTFHADICTLSEKERQIKQTALVE 531
Db 496 GVGQCCTSSYANRRPCFSSLVVDYVYPPAFSDDKDFIHKDLCAQGVQALQTMKQBEFLN 555
Qy 532 LVKHKPKATKQKLVKAVMDMDFAAFVEKCKADKDETCFAEEGKKLVAASQAALGL 585
Db 556 LVKQKQITEQLTEVIADFSGLLEKCKCQGEQVEVCAEEGKKLISKTRALGV 609
```

RESULT 29

```
Q8MJ76
ID Q8MJ76 PRELIMINARY; PRT; 609 AA.
AC Q8MJ76
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Alpha-fetoprotein.
GN Name=afp;
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_Taxid=9615;
RN [1]
RP SEQUENCE FROM N.A.
```

```
RA Furuichi M., Neo S., Hisasue M., Tsuchiya R., Watanabe M.,
RA Hashizaki K., Hisamatsu S., Yamada T.,
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB089789; BAC07513.1; -.
DR HSP; P02768; IHK2.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR001703; Alpha-fetoprot.
DR Pfam; PF00273; Serum_albumin; 3.
DR PRINTS; PR00803; AFEPTOPROTEIN.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 609 AA; 68782 MW; BE4B8250C5AF2AF0 CRC64;

Query Match 32.8%; Score 1242.8; DB 2; Length 609;
Best Local Similarity 40.5%; Pred. No. 6.9e-23;
Matches 231; Conservative 113; Mismatches 225; Indels 2; Gaps 2;

Qy 16 EENFKALVLIAPFAVQLQCCPEDHVKLVNVEVTEFAKTCTVADESAENCDKSLHTLFGDKLC 75
Db 40 EMLVLDLATIFFAQFVQEAITYEYKSVKVKDILTVIEKSTGSEQPGGCLNQLPAFLSEIC 99
Qy 76 TVATLRETYGEMADCCAKQEPNERNECFLOHKD-DNPNLPLRVREVDVMTAFHDNEETF 134
Db 100 HEKEISEKYG-LADCCSQSEERHNCFLAHKKAAPPISPPFVABPVTSCAYEENRDMF 158
Qy 135 LKYLVIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLLPKLDLDEDEGKASSA 194
Db 159 MNRVYIYIARRHPFLYAPTILSLAAHYGKILPICKKAENAVECFQYTSILITKELRESSL 218
Qy 195 KORLKCAQLQKFGERAFKAWAVARLSORFPKABFAEVSKLVTDLTKVHTECHGDLLECA 254
Db 219 LNQHICAVMRNFGPRTTFRATVTKLQSKFSKANFTETQKLVLDVAHTEBCCRGVLECL 278
Qy 255 DRRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKD 314
Db 279 QDGEKINSYICSQDILSSKIADCCCKLPLELQGCIIHAENDGKPEGLSPNLNRFLBERD 338
Qy 315 VCKNYAEAKVDFGLMFLYEHARRHPDYSVVLLRLAKYETTLKCCAAADPHECYAKVF 374
Db 339 FNGFSSREKDLFWARFYEYSRRHTKLAVPVLVAKGYQELLEKCKSQSENFLECDQKE 398
Qy 375 DEFKPLVEEPQNLIKONCELFQELGEYKFNALLVRYTKKVPQVSTPTLVEVSRLNLGKVG 434
Db 399 EELEKYIQESQALAKRSCGLFQKLGYYLQNAFLVAYTKKAPQLTPPELMAFTRKMTAA 458
Qy 435 SKCKHPKAPKMPCAEDYLSVVLNQLCVLHKEKTPVSDRYTKCTTESLVNRRPCFSALEVD 494
Db 459 ATCCQLSEDLQACGEGAADLIIGQLCIRHEETPINPGVQCSCSSYANRRPCFSSLVVD 518
Qy 495 ETYVPEKPEFNAETFTFHADICTLSEKERQIKQTALVELVKHKPKATKQKLVKAVMDMDFAAF 554
Db 519 ETYVPSFSDKDFIHKDLCAQGVQALQTMKQBEFLNVLVKQKQITEQLTEVIADFSGL 578
Qy 555 VEKCKKADDETCFAEEGKKLVAASQAALGL 585
Db 579 LEKCCQGEQVEVCAEEGKKLISKTRALGV 609

RESULT 30
Q8MJ76
ID Q8MJ76 PRELIMINARY; PRT; 610 AA.
AC Q8MJ76
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Alpha-fetoprotein.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```



**THIS PAGE BLANK (USPTO)**